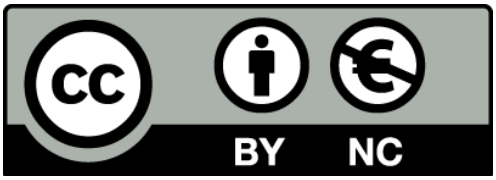
Flavour Tagging developments within the LHCb experiment

Marc Grabalosa Gándara



UNIVERSITAT DE BARCELONA

Department d'Estructura i Constituents de la Matèria

Flavour Tagging developments within the LHCb experiment

by

Marc Grabalosa Gándara

Flavour Tagging developments within the LHCb experiment

Memòria de la tesis presentada per

Marc Grabalosa Gándara

per optar al títol de Doctor en Física

Director de tesi:

Marco Musy

Tutor:

Lluís Garrido Beltrán

Programa de doctorat en Física de l'EEES

Department d'Estructura i Constituents de la Matèria

Barcelona, Març 2012

"Quan algú comprèn que obeir lleis injustes és contrari a la seva dignitat d'home, cap tirania no el pot dominar."

Mohandas K. Gandhi

UNIVERSITAT DE BARCELONA

Flavour Tagging developments within the LHCb experiment

by Marc Grabalosa Gándara

Facultat de Física
Department d'Estructura i Constituents de la Matèria

Abstract

Flavour Tagging at the LHCb experiment is a fundamental tool for the measurement of B oscillations and the study of CP violation. This document explains the development of different tagging techniques and the different strategies used to combine them to determine the flavour of the B meson as precisely as possible. The response of the tagging algorithms also needs to be optimized and calibrated. Both procedures are described using the available LHCb datasets corresponding to various integrated luminosities. First results on the tagging performances are shown for different control channels and physics measurements.

Agraïments

Durant tots aquests anys de doctorat, hi ha hagut molta gent que m'ha donat el seu suport i m'ha ajudat molt. Per això no vull acabar aquesta tesi sense haver-los donat les gràcies.

Primer de tot, vull donar les gràcies al meu director i al meu tutor, en Marco Musy i en Lluís Garrido per oferir-me l'oportunitat de fer la tesi en el grup d'altes energies de la UB i per poder participar en la "LHCb Collaboration". En segon lloc, m'agradaria expressar la meva més sincera gratitud a la Míriam, que ha tingut la paciència de llegir-se tota la tesi i fer-me comentaris molt útils que de ben segur han ajudat a que la tesi agafés la forma que té ara.

També m'agradaria agrair l'ajut que m'han ofert quan era necessari altres persones vinculades a aquest grup: l'Eugeni, en Ricardo, l'Hugo, l'Alessandro, en Jordi, l'Elies, l'Antonio, en Cédric, l'Albert, en Ricci, en Xavier, l'Adrià, l'Edu, en David, l'Albert, el Vicente... A tots ells, moltes mercès. I vull agrair també el suport que m'han donat els meus companys de despatx o de facultat, no vinculats al grup, amb qui també he compartit molts moments: discusions, xerrades, etc. Ells són en Miguel, en Clemens, en Jaume, en Juan, l'Aldo i l'Ari.

The tagging group of the LHCb has had, of course, a very important role for my thesis and I also want to mention them: Sara, Antonio, Denis, Emilie, Olivier, Marta and specially Stefania. They have also helped me a lot during these last years. To all of them: Many thanks for your support!

My feelings of gratitude would also like to travel to reach all the people who I met at CERN, Javi, Vicente, Neus, Snezana, Beni, and more. Et aussi a mes amies de Genève, Christina et Anna.

I per acabar, vull adreçar una abraçada molt especial als meus amics Arnau i Pep per tota la calor d'aquests anys, i a en Niko, l'Ester i la Mireia. Moltes mercès també per aguantar-me amb el vostre amor, als meus pares i a la meva germana: Gràcies! Us estimo.

Però, sobretot, a qui més vull agrair el seu suport incondicional és a la Laia, sense ella, tot això no hauria estat possible. T'estimo. A tu i a la meva nova alegria de viure, la meva fada d'aigua.

Contents

Resum

Introducció

En aquesta tesi presento detalladament l'etiquetatge de sabor (*Flavour Tagging*) a l'LHCb. L'etiquetatge de sabor és una eina bàsica per poder realitzar mesures d'oscil·lacions de mesons B i de violació de CP (Càrrega Paritat), processos indispensables per a la recerca de nova física més enllà del Model Estàndard.

Física de l'*LHCb*

El Model Estàndard (ME), introduit el 1961 per Glashow, Weinberg i Salam [1–3], és el model vigent que s'utilitza per descriure les partícules fonamentals i les interaccions entre elles. Aquest model s'ha anat consolidant durant la segona meitat del segle XX amb moltes evidències experimentals i prediccions teòriques. Tot i això, té alguns punts febles i tot sembla indicar que el ME és només una teoria efectiva d'una teoria més general. Un exemple d'això són les discrepàncies entre les mesures cosmològiques i les predites pel ME de la violació de CP, i per això aquesta simetria juga un rol tant important en la recerca de nova física. Els experiments de l'LHC són el lloc adient per a buscar-ne els efectes.

Model Estandard

Segons el ME, hi ha dos tipus de partícules fonamentals: els fermions, que constitueixen la matèria; i els bosons, considerats els portadors de les forces. Tota la matèria de l'univers està composta de dos tipus de fermions: quarks i leptons, que s'agrupen en tres famílies amb propietats idèntiques a excepció de la massa. Els quarks estan confinats en estats lligats anomenats hadrons, que poden ser: mesons, quan estan formats per quarks i antiquarks ($q\bar{q}$); i barions, formats per tres quarks o tres antiquarks (qqq, $\bar{q}\bar{q}\bar{q}$). Els leptons poden ser carregats (e, μ, τ) o bé neutres (ν_e, ν_μ, ν_τ).

Les interaccions descrites pel ME inclouen la força electromagnètica, la força nuclear dèbil i la força nuclear forta i es poden explicar mitjançant l'intercanvi de partícules mediadores, els bosons. La força nuclear forta es descriu mitjançant la cromodinàmica quàntica (QCD), en la qual les partícules portadores de la interacció són els gluons. L'electromagnetisme i la interacció dèbil es descriuen mitjançant la teoria electrofeble, en la qual els portadors són els bosons $W^{\pm}$ i Z. En el model aquests bosons adquireixen massa a través del model de Higgs de ruptura

espontània de simetria. Aquest model explica també l'origen de les masses dels fermions i de la barreja entre quarks de diferents families a través de les corrents carregades que vénen descrites per la matriu Cabibbo-Kobayashi-Maskawa (CKM).

La matriu CKM és la font de la violació de CP del ME [14]. Aquesta té lloc quan no hi ha una simetria sota una conjugació de càrrega i paritat (CP). Va ser observada per primera vegada el 1964 per Cronin i Fitch en els mesons K neutres a un nivell del 0.2% [16], i s'ha observat també en mesons B neutres a les factories de mesons B com BaBar o Belle [17, 18].

La matriu CKM relaciona els estats propis de massa dels quarks amb els estats propis de la interacció dèbil. Aquesta matriu pot ser descrita per 4 paràmetres: tres angles de rotació i una fase complexa. En la parametrització de Wolfenstein aquesta es pot escriure com a:

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \quad (1)$$

La unitarietat de la matriu CKM implica 6 relacions d'ortogonalitat entre els seus elements. Aquestes poden ser representades al pla complexe com a triangles, per això reben el nom de "Triangles Unitaris". La mesura dels angles interns dels triangles ens permeten testejar la matriu CKM i mesurar els efectes de la violació de CP entre elements de la matriu CKM. El triangle representat en la Figura 1 s'anomena "Triangle Unitari" i representa la relació d'unitarietat més accessible. Desviacions de la forma triangular és un indici de nova física més enllà del ME.

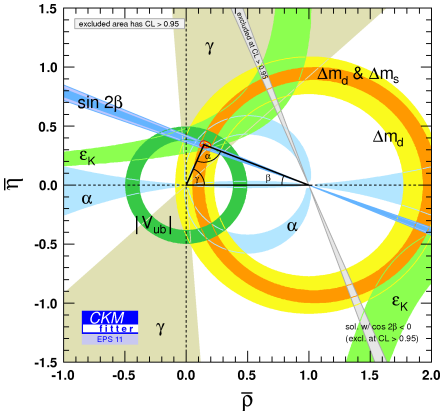


FIGURE 1: Delimitació del "Triangle Unitari" feta a partir de diferents mesures (Taula 1.1).

Mesons B

Els mesons B són partícules formades per un quark b o un antiquark $\bar{b}$ i un altre quark. La interacció feble permet un fenomen anomenat oscil·lacions de mesons B, a través del qual, un

mesó B^0 pot convertir-se en un $\bar{B}^0$ i viceversa. Aquest procés va ser observat per primer cop el 1987 [33, 34] i pot ser descrit mitjançant diagrames de caixa en la interacció feble, tal i com es mostra en la Figura 2.

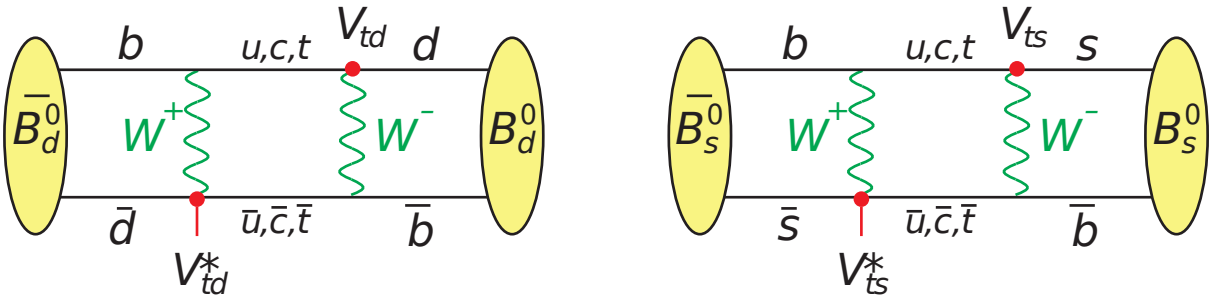


FIGURE 2: Diagrama de caixa per les oscilacions $B^0\bar{B}^0$, per mesons B_d i B_s.

L'evolució dels mesons B neutres, i la seva oscil·lació, pot ser descrita mitjançant el Hamiltonià de la interacció feble, el qual depèn de la diferència de masses i del temps de vida dels estats propis del Hamiltonià ($\Delta m_d = m_H - m_L$, $\Delta\Gamma = \Gamma_H - \Gamma_L$). Les mesures prèvies abans de l'entrada en funcionament del LHCb eren: $\Delta m_d = 0.507 \pm 0.005$ ps^{-1} [6], i pels mesons B_s: $\Delta m_s = 17.77 \pm 0.10 \pm 0.07$ ps^{-1} [6]. L'LHCb en el seu primer any de funcionament ha mesurat Δm_d i Δm_s amb un error similar: $\Delta m_d = 0.499 \pm 0.032(\text{stat}) \pm 0.003(\text{syst})$ ps^{-1} [36] i $\Delta m_s = 17.63 \pm 0.11(\text{stat}) \pm 0.02(\text{syst})$ ps^{-1} [37].

Violació de CP

La violació de CP correspon a diferències entre processos físics i els seus processos conjugats via CP. Com que els estats congujats estan relacionats via conjugats hermítics, la presència d'acoplaments complexes provinents de la matriu CKM al Hamiltonià pot trencar la simetria sota transformacions de CP. Per tal de quantificar la violació de CP és útil identificar quantitats independents de les convencions de fase i poder mesurar asimetries entre un procés i el seu conjugat. La violació de CP pot ser de tres tipus:

- **Violació de CP directa o violació de CP en el decaïment.** Es dóna en la desintegració de mesons carregats o neutres quan l'amplitud de probabilitat d'una desintegració ($B \to f$) i de la seva conjugada ($\overline{B} \to \overline{f}$) són diferents. En aquest cas, és útil mesurar:

$$\mathcal{A}_{CP} = \frac{\Gamma(\overline{B} \to \overline{f}) - \Gamma(B \to f)}{\Gamma(\overline{B} \to \overline{f}) + \Gamma(B \to f)} \tag{2}$$

- **Violació de CP indirecta o violació de CP en la barreja.** Succeeix en mesons neutres que són estats propis de la matriu de masses però no de la transformació de CP.

És un fenòmen estretament relacionat amb la barreja, i respon a diferències entre el pas de $B^0 \to \bar{B}^0$ i del pas invers $\bar{B}^0 \to B^0$. L'asimetria observable en aquest cas és:

$$\mathcal{A}_T = \frac{\Gamma(|\bar{B}^0(t)\rangle \to f) - \Gamma(|B^0(t)\rangle \to \overline{f})}{\Gamma(|\bar{B}^0(t)\rangle \to f) + \Gamma(|B^0(t)\rangle \to \overline{f})} \tag{3}$$

- **Violació de CP en la interferència de la barreja i del decaïment.** Es dóna quan les desintegracions a estats finals propis de CP són accessibles als dos mesons neutres. En aquest cas podem mesurar l'asimetria:

$$\mathcal{A}_{f_{CP}}(t) = \frac{\Gamma(|\bar{B}^0(t)\rangle \to f_{CP}) - \Gamma(|B^0(t)\rangle \to f_{CP})}{\Gamma(|\bar{B}^0(t)\rangle \to f_{CP}) + \Gamma(|B^0(t)\rangle \to f_{CP})} \tag{4}$$

L'experiment LHCb

El Gran Colisionador d'hadrons

El gran colisionador d'hadrons (LHC en les seves sigles en anglès) és l'accelerador de partícules més potent i més gran del món. Està situat al CERN, a la regió fronterera entre Suïssa i França, al costat de la ciutat de Ginebra. Ocupa un túnel de 27 km en forma d'anell a 100 m de profunditat i està compost per dos acceleradors paral·lels pels quals circulen protons a velocitats properes a les de la llum, en sentits oposats. L'LHC està dissenyat per fer col·lidir els protons a una energia nominal en el centre de masses de 14 TeV en quatre punts diferents, on hi ha quatre detectors que analitzen les col·lisions. Aquests detectors són: $ATLAS$ i CMS, detectors amb caràcter general que busquen noves partícules com el Higgs; $ALICE$, que estudia un nou estat de la matèria: el plasma de quarks i gluons; i l'$LHCb$, que està dedicat a la física del sabor, en particular a la dels mesons B.

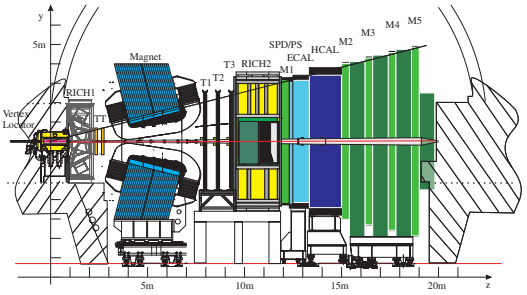


FIGURE 3: Esquerra: Esquema de l'LHC, amb els detectors ATLAS, CMS, LHCb i ALICE. Dreta: vista esquemàtica de l'experiment LHCb juntament amb tots els seus subdetectors.

LHCb

L'experiment LHCb va ser concebut per estudiar la física de les desintegracions de mesons B produïdes en les interaccions protó-protó a 14 TeV amb una freqüència d'interacció de 40 MHz a l'accelerador LHC. L'acceptància de l'LHCb comprèn angles propers a l'eix d'incidència, i es restringueix a un sol costat de la regió d'interacció. Això està motivat per la distribució dels mesons B que es produeixen bàsicament en parelles i en la mateixa direcció del feix. Dins d'aquesta acceptància es preveu la producció d'aproximadament 10^{12} $b\bar{b}$ per any nominal (2 fb^{-1}).

LHCb està constituit per diferents subdetectors, ordenats segons la distància a la regió d'interacció són:

- VELO (*Vertex Locator*): localitza els vèrtexs i en separa els primaris (de la col·lisió) dels secundaris (de les desintegracions de mesons B).
- RICH: és un detector d'anells de radiació produits per la radiació Cherenkov que es produeixen quan una partícula circula per un medi a una velocitat més gran que la de la llum en aquest medi. S'utilitza per identificar partícules.
- Un Sistema de detectors de traces, format pel TT, IT i l'OT: s'encarrega de mesurar la trajectòria de les partícules carregades sota l'efecte d'un camp magnètic.
- Els calorímetres: formats per l'SPD, el PS, un calorímetre electromagnètic (ECAL) i un d'hadrònic (HCAL) que s'utilitzen per mesurar la posició de les cascades electromagnètiques i hadròniques. També s'usa en el primer nivell de filtratge i en la detecció de partícules neutres.
- Les cambres de muons: s'utilitzen per mesurar el moment transvers dels muons.

La producció de mesons B es pot distinguir per la presència de vèrtex secundaris i d'un alt moment transvers en els productes de les desintegracions. Tot i això, és necessari un sistema de filtratge, o trigger, per reduir el ritme d'esdeveniments dels 40 MHz als 2 kHz que són els permessos per la capacitat d'emmagatzematge. Aquest filtratge consta d'un primer nivell, anomenat *Level-0*, completament integrat en l'electrònica del detector, que rebaixa el ritme d'esdeveniments fins a 1 MHz. El segon nivell, anomenat *High Level Trigger* (HLT) aconsegueix la reducció restant mitjançant l'ús d'algoritmes operant en una granja de computació. La presa de dades i el control del detector s'efectua mitjançant: el sistema d'adquisició de dades (DAQ), el sistema de mesura de temps i de control ràpid (TFC) i el sistema de control de l'experiment (ECS).

Durant el primer període de presa de dades, l'energia dels feixos era de 3.5 TeV per cadascun d'ells. El 2010 es van recollir ~ 35 pb^{-1}, que es van poder utilitzar per posar en evidència el bon

funcionament del detector i per realitzar algunes mesures. A l'estiu del 2011, ja s'havien recollit ~ 370 pb^{-1} i realitzat moltes mesures en el camp de les oscil·lacions de B o en el camp de la violació de CP. Algunes d'aquestes mesures van donar els resultats més precisos que s'havien aconseguit fins al moment. A finals del 2011, la lluminositat recollida ja ha arribat a ~ 1 fb^{-1} i es preparen molts anàlisis per obtenir noves mesures.

Etiquetatge de sabor

En les mesures de violació de CP i per les oscil·lacions de mesons B és necessari conèixer el sabor de l'estat inicial del mesó B reconstruit, per saber si contenia un quark b o un antiquark $\bar{b}$. Aquest procés s'anomena etiquetatge de sabor (*Flavour Tagging*) i a l'LHCb es realitza per mitjà d'uns algoritmes que estudien la càrrega dels productes de decaïment de l'altre mesó B (B_{opp}) del succés, o bé mitjançant la identificació de pions o kaons produïts en el procés de fragmentació del mesó B reconstruit o de senyal (B_{sig}). Reben el nom de *OS (opposite side) taggers* en el primer cas i de *SS (same side) taggers* en el segon. A més, aquests algoritmes estimen també la probabilitat de fallar la decisió. En la Figura 4 es pot veure de forma esquemàtica un succés $b\bar{b}$ i els diferents algoritmes que es poden utilitzar per a l'etiquetatge de sabor.

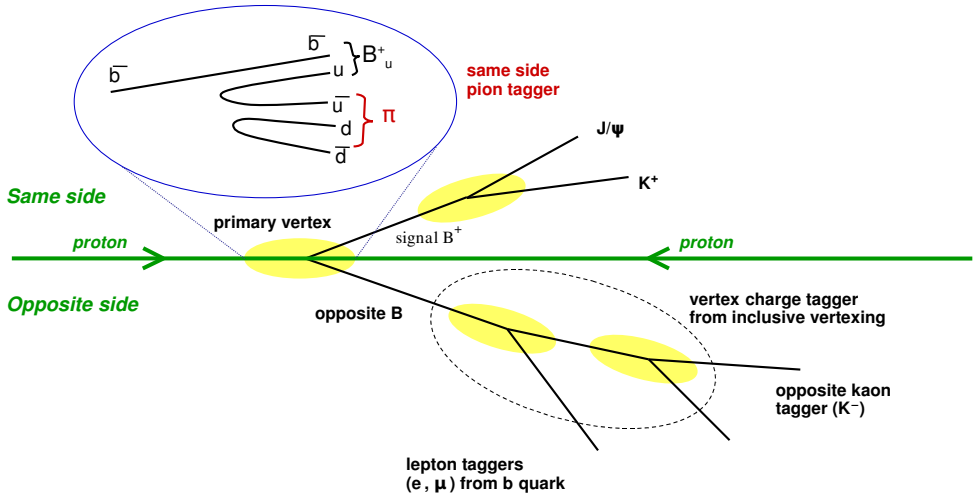


FIGURE 4: Visió esquemàtica d'un succés $b\bar{b}$. L'anomenat "mateix costat" (SS) correspon a la part del decaïment de la B reconstruida, amb pions o kaons com a partícula d'etiquetatge; el "costat oposat" (OS) correspon a l'altre B produïda en la col·lisió, amb leptons i kaons com a partícules d'etiquetatge juntament amb la reconstrucció del vèrtex secundari.

El tant per cent dels esdeveniments pels quals es troba una resposta s'anomena eficiència (ϵ_{tag}). A més, l'etiquetatge tampoc és perfecte i hi ha una certa probabilitat d'obtenir el resultat equivocat. Aquesta probabilitat d'errar s'anomena ω. La conseqüència d'aquests etiquetatges erronis és que l'asimetria es veurà diluida per un factor $(1 - 2\omega)$ respecte a la asimetria real. Mitjançant aquests dos paràmetres es pot calcular l'eficiència efectiva (ϵ_{eff}), que està directament

relacionada amb la sensitivitat de l'asimetria mesurada i indica la degradació estadística de la mesura deguda a ω. L'eficiència i la probabilitat d'equivocar-se poden determinar-se directament de les dades. En especial, el valor de ω es pot determinar directament mitjançant els canals de control, que són canals que decauen a estats de sabor específics i no presenten violació de CP, com el $B^+ \to J/\psi K^+$ o el $B^0 \to J/\psi K^*$. En canvi, la mesura d'ω no és possible en els canals de CP, com el $B^0 \to J/\psi K_s$, en el qual s'ha de confiar en la mesura obtinguda a través dels algoritmes d'etiquetatge.

Si R, W i U són el nombre de successos ben etiquetats, mal etiquetats i no etiquetats respectivament, es poden definir les variables de la forma següent:

$$\epsilon_{tag} = \frac{R+W}{R+W+U} \qquad \omega = \frac{W}{R+W} \qquad \epsilon_{\mathit{eff}} = \epsilon_{tag}(1-2\omega)^2 \tag{5}$$

No totes les partícules d'un succés poden ser utilitzades per l'etiquetatge dels quarks. S'usa primer una preselecció per tal de maximitzar la probabilitat que les partícules provinguin del mesó B oposat o siguin pions i kaons produits en la fragmentació de la B reconstruïda. Aquesta preselecció està resumida en la Taula 3.1.

Algoritmes d'etiquetatge del costat oposat

S'utilitzen per endevinar el sabor del quark de la B de senyal (B_{sig}) mirant el sabor de l'altre B de l'esdeveniment (B_{opp}). Per això, s'observa la càrrega del leptó provinent de desintegracions semileptòniques o bé la càrrega del kaó provinent de decaïments de $b \to c \to s$. També s'usa la càrrega del vèrtex secundari, reconstruit a partir dels seus productes de decaïment. En el cas que el mesó oposat sigui neutre ($\sim 50\%$) una dilució vindrà donada per la probabilitat d'aquest mesó d'oscil·lar. Els algoritmes d'etiquetatge, o taggers, del costat oposat són:

- **Muó**. Busca muons provinents de desintegracions semileptòniques del B_{opp}. Assumint que la B no oscil·la, la càrrega del muó identifica unívocament el sabor del quark b. Per a una bona identificació dels muons s'apliquen talls en $\Delta LL_{\mu-\pi}$, $\chi^2/d.o.f$, p_T i s'usen algoritmes extra per tal de no seleccionar clons ni duplicats.

- **Electró**. Busca electrons de les desintegracions semileptòniques del B_{opp}. Per maximitzar l'eficiència efectiva de l'etiquetatge, les particules són identificades com a electrons si passen cert talls en $\Delta LL_{e-\pi}$ i p_T. A més, han d'estar en l'acceptància del calorímetre i que la fracció energia-moment, així com la càrrega dipositada en el VELO, siguin superiors i inferiors a certs valors respectivament.

- **Kaó**. S'utilitza el kaó produit en el decaïment $b \to c \to s$. Per tal de maximitzar la probabilitat de seleccionar aquest kaó s'utilitzen talls en l'identificació de les partícules,

particularment en p_T, $\chi^2/d.o.f$, així com en el paràmetre d'impacte respecte al vèrtex primari.

- **Vèrtex secundari**. Està basat en la reconstrucció inclusiva de vèrtex secundaris corresponents al B_{opp}. Utilitza la càrrega del vèrtex per determinar el sabor del quark.

  El procés comença construint parelles de traces que es puguin utilitzar com a llavor, per tal de construir el vèrtex. S'apliquen certs talls a les traces i a la llavor per maximitzar la probabilitat que aquesta estigui formada per dues traces provinents del B_{opp}. Bàsicament són talls en paràmetre d'impacte, p_T, $\chi^2/d.o.f$ i una *likelihood* construïda a partir de certes variables de la llavor. Com a origen del vèrtex secundari sempre es selecciona la llavor amb la *likelihood* més alta. Un cop s'ha triat la llavor s'afegueixen traces al vèrtex si aquestes passen certs talls dissenyats per maximitzar la probabilitat de que les traces provinguin del B_{opp}.

  Un cop reconstruit tot el vèrtex, se'n mesura la càrrega i es pesa en funció del p_T. En aquest punt, s'apliquen certs talls per tal de maximitzar l'eficiència efectiva, com per exemple, talls en la càrrega del vertex, en la suma del p o del p_T de totes les traces del vèrtex.

Totes les variables estàn detallades en el capítol 3. I en les Taules 4.17 i 4.19 hi ha els talls utilitzats en dades reals que maximitzen l'eficiència efectiva o el poder del tagging.

Algoritmes d'etiquetatge del mateix costat

Aquests algoritmes determinen directament el sabor del quark B utilitzant la correlació en la cadena de fragmentació. Així, quan un B^0 es reconstrueix, un quark $\bar{d}$ extra queda disponible per formar un pió (π). En el cas d'un B_s^0, un quark $\bar{s}$ extra permet formar kaons carregats en el 50% dels casos. I en el cas del B^+, un quark $\bar{u}$ ens permet formar pions o kaons. Els algoritmes del mateix costat identifiquen la càrrega d'aquests pions o kaons per determinar el sabor del quark b.

- **SS** π. Per tal d'assegurar-se una bona identificació de pions provinents del vèrtex primari i de la cadena d'hadronització de la B de senyal, s'apliquen talls en p_T, p, $\Delta LL_{k-\pi}$, $\Delta LL_{p-\pi}$, parametre d'impacte, així com també es calcula la diferència de masses (ΔQ) formada pel sistema $B^0\pi$ i del mesó B reconstruit.

- **SS K**. De la mateixa manera que pel π, s'apliquen talls en p, p_T, parametre d'impacte, $\chi^2/d.o.f$, $\Delta LL_{k-\pi}$, ΔLL_{k-p}, ΔQ, així com també en les diferències en pseudo-rapiditat i angle ϕ entre el kaó i el mesó B reconstruit.

Igual que pels algoritmes del costat oposat, totes les variables estan detallades en el capítol 3, i en les Taules 4.10 i 4.18 hi ha els talls utilitzats amb dades reals que maximitzen l'eficiència efectiva o poder del tagging pel algoritmes del mateix costat.

Altres algoritmes

Durant la fase de disseny dels algoritmes, van sorgir algunes idees per tal de millorar-los. La primera era la introducció d'un segon conjunt de talls per tal de seleccionar electrons i muons provinents del decaïment $b \to c \to l$ que tenen la càrrega oposada, i que permet millorar un $\sim 20\%$ aquests dos algoritmes. També es va dissenyar un algoritme que buscava un kaó provinent de la fragmentació del B_{opp}, que indicaria inequívocament el sabor del quark b. El problema és que els successos amb B_s són pocs, i l'eficiència de la reconstrucció, a més, seria molt baixa, ja que requereix també l'existència d'un vèrtex secundari. El mètode més exitós va ser el disseny d'un algoritme per identificar D^0 en el costat oposat, ja que mirant-hi els productes de decaïment es pot determinar el sabor del quark b. La idea consisteix en buscar parelles de pions i kaons al voltant de la massa de la D amb uns certs talls que maximitzin l'eficiència efectiva. Finalment es va poder crear un algoritme amb una eficiència del $\sim 0.3\%$ però no es va arribar a introduir a causa de l'alta correlació amb l'algoritme que reconstrueix kaons i el que reconstrueix el vèrtex secundari.

Decisió d'etiquetatge i probabilitat d'etiquetar malament

La desició del sabor del quark d ve donada pels diferents algoritmes d'etiquetatge i serà $d = +1$ si l'hadró conté un quark $\bar{b}$ i $d = -1$, si conté un quark b. La probabilitat del *tagger* de fallar la desició (ω) s'estima mitjançant xarxes neuronals entrenades amb simulacions Monte Carlo (MC) mitjançant el paquet *MultiLayer Perceptron* (MLP) de root [82]. Aquesta ω és clau per a la combinació dels algoritmes d'etiquetatge que permeten una millor estimació tant del sabor del quark com de la probabilitat d'encert del *tagger*, cosa que és necessària per a la majoria d'anàlisis de violació de CP.

Les xarxes neuronals estan formades per una sola capa intermitja i depenen de variables geomètriques i cinemàtiques de la partícula d'etiquetatge, així com del nombre de vèrtex reconstruits, el p_T de la B_{sig} o de la multiplicitat de les traces. A partir de la resposta de la xarxa neuronal (net_i) es pot estimar fàcilment la probabilitat d'etiquetar malament, p.e $\omega_{vtx} = a_{vtx} + b_{vtx} net_{vtx}$. Si la parametrització és correcte es podrà observar que la ω mesurada correspon completament a la ω estimada pels algoritmes d'etiquetatge, anomenada η, tal i com es pot veure en la Figura 5.

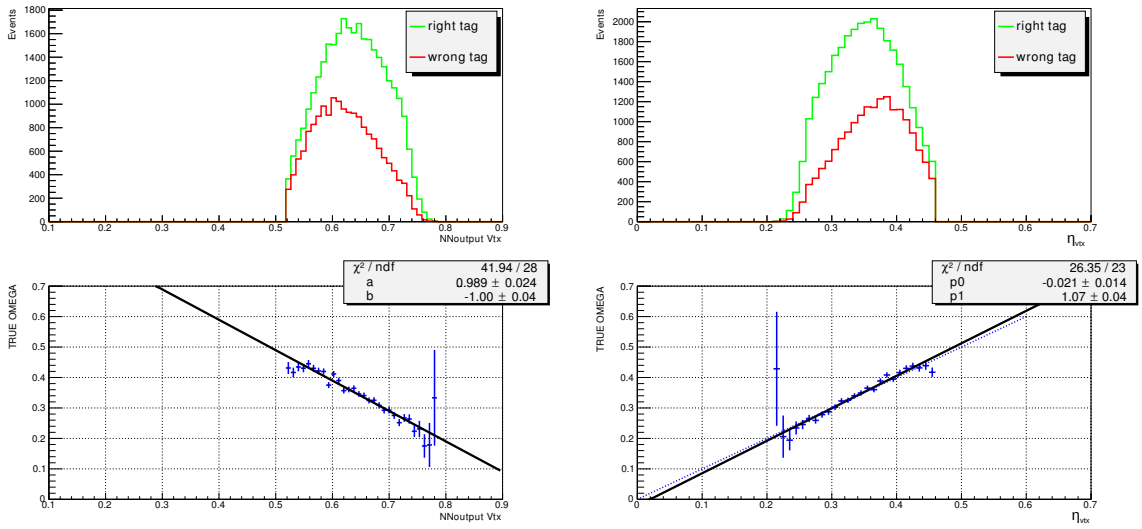


FIGURE 5: OS vèrtex secundari per simulacions MC d'esdeveniments $B^+ \to J/\psi K^+$. El verd correspon a esdeveniments ben etiquetats, i el vermell als mal etiquetats. A la segona fila es pot veure, a l'esquerra, el valor d'η directament extret de la xarxa neuronal (NNoutput) i a la dreta, la ω mesurada en funció de la η estimada (η) per l'algoritme d'etiquetatge.

Optimització i calibració dels algoritmes d'etiquetatge

Els algoritmes d'etiquetatge van ser desenvolupats inicialment basant-se en simulacions MC. Però a causa de les diferències entre el MC i les dades reals o entre les diferents reconstruccions de les dades reals, pot ser que els algoritmes d'etiquetatge no siguin òptims. És necessari, doncs, un procés d'optimització i calibració per tal d'obtenir una ω creïble i maximitzar els resultats de l'etiquetatge (ϵ_{eff}) en dades. Per a aquests processos s'utilitzen canals de control on la ω pot ser fàcilment mesurada comparant el sabor real del quark amb aquell predit pels algoritmes d'etiquetatge.

L'optimització busca el conjut de talls en les variables usades per seleccionar el candidat a esdevenir *tagger* o particula d'etiquetatge que maximitzi els resultats o ϵ_{eff}, mentre que la calibració corregeix l'estimació de la fracció d'esdeveniments on el tagging dóna una mala resposta (η) mitjançant una relació lineal amb la ω que es pot mesurar en els canals de control de la forma següent:

$$\omega = p_0 + p_1 \cdot (\eta - < \eta >) \tag{6}$$

Per tal d'adaptar l'etiquetatge de sabor a les dades reals, l'optimització i calibració s'han anat refent a mesura que s'anaven tenint noves dades per a 35 pb^{-1}, 370 pb^{-1} i 1 fb^{-1}.

Optimització i calibració amb 35 pb^{-1}

La primera optimització utilitzava el canal de control amb més esdeveniments recollits, el $B^0 \to D^{*-}\mu^+\nu_\mu$. En aquest cas, l'$\omega$ es pot obtenir mitjançant un ajust a la oscil·lació de sabor, a través dels esdeveniments que han canviat de sabor (mix) i els que no (unmix) en funció del temps de decaïment (t). Aquesta oscil·lació, en absència de soroll, pot ser interpretada de la forma següent:

$$A(t) = \frac{N^{\mathrm{unmix}}(t) - N^{\mathrm{mix}}(t)}{N^{\mathrm{unmix}}(t) + N^{\mathrm{mix}}(t)} = (1 - 2\omega)\cos(\Delta m_d t) \tag{7}$$

Per tal de seleccionar el mínim soroll possible només s'usaven els sedeveniments situats en les finestres de massa del $D^* - D^0$ ($143.7\ \mathrm{MeV/c^2} < m(K\pi\pi) - m(K\pi) < 147.8\ \mathrm{MeV/c^2}$), del D^0 ($|m(K\pi) - M_{D^0}| < 24\ \mathrm{MeV/c^2}$) i del D^* ($|m(K\pi\pi) - M_{D^*}| < 24\ \mathrm{MeV/c^2}$). També es descartaven els esdeveniments amb $t < 0.2$ ps per descartar soroll provinent del vèrtex primari (prompt). Uns $\sim 48\,000$ esdeveniments van permetre optimitzar l'etiquetatge de sabor o ϵ_{eff} mitjançant aquest mètode. Pel SSK, a causa de les poques dades recollides, el *tagger* només es va poder optimitzar mitjançant decaiments $D_s^+ \to \phi\pi^+$.

Per a la calibració es van fer servir $\sim 11\,000$ esdeveniments del canal $B^+ \to J/\psi K^+$. La distribució d'η es pot extreure d'una finestra de $\pm 40\ \mathrm{MeV/c^2}$ al voltant de la massa del B^+, i comparant-la amb la ω mesurada, s'en pot extreure una correcció.

Optimització i calibració amb 370 pb^{-1}

Quan es van recollir dades suficients, el procés d'optimització es va realitzar mitjançant el canal $B^+ \to J/\psi K^+$. En aquest cas, la ω es pot obtenir comparant la decisió dels algoritmes d'etiquetatge amb el sabor de la B de senyal. Els talls trobats eren molt similars als trobats amb el canal $B^0 \to D^{*-}\mu^+\nu_\mu$ i això confirmava que el funcionament dels algoritmes d'etiquetatge era molt proper a l'òptim. A més, es va intentar trobar un conjunt de talls en funció del nombre de vèrtex primaris o de la multiplicitat de l'event. Els talls òptims però, eren els mateixos, independentment de la situació inicial. La calibració es va realitzar de la mateixa forma que amb 35 pb^{-1}, aquest cop amb la novetat que, pel SSπ, un polinomi de segon grau va ser necessari per corregir la η.

Optimització i calibració amb 1 fb^{-1}

La optimització es va realitzar mitjançant esdeveniments $B^+ \to J/\psi K^+$ pels OS taggers, amb $B^0 \to D^{*-}\mu^+\nu_\mu$ pel SSπ i amb $B_s^0 \to D_s^-\pi^+$ pel SSK. Solament es van aplicar petites correccions en els taggers respecte als talls trobats anteriorment. D'altra banda, es va provar una optimització amb el canal $B^0 \to D^-\pi^+$ tal i com es va fer amb el $B^0 \to D^{*-}\mu^+\nu_\mu$. Per a la calibració dels OS *taggers*, es va tornar a utilitzar el canal $B^+ \to J/\psi K^+$. A la Figura 6 es poden veure els gràfics de la calibració del OSK per soroll i senyal. En el cas del SSπ es va utilitzar el $B^0 \to J/\psi K^*$, i pel SSK, es va poder fer servir el canal $B_s^0 \to D_s^-\pi^+$ per optimitzar i calibrar el tagger mitjançant l'ajust a l'oscil·lació de mesons B_s.

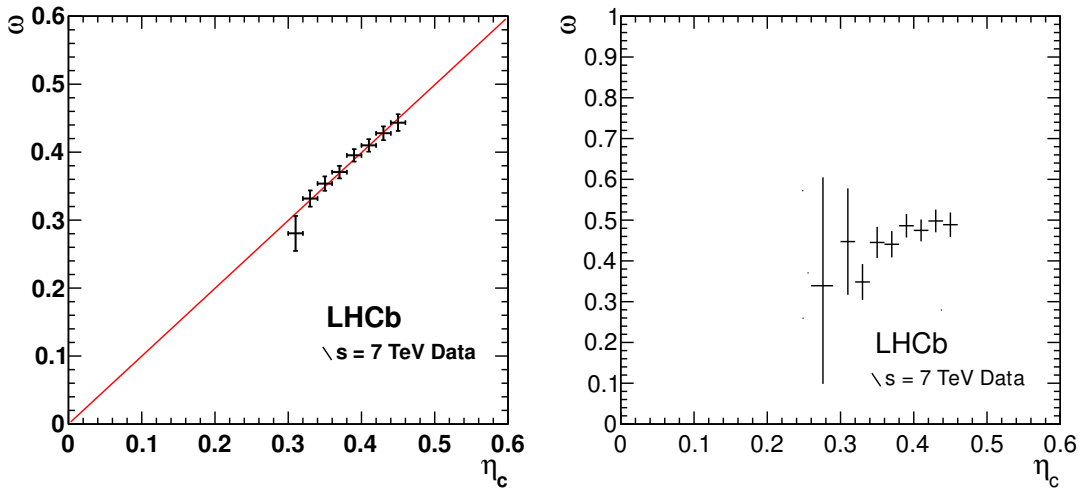


FIGURE 6: Plots de calibració per l'OSK *tagger* per a esdeveniments de senyal (esquerra) i de soroll (dreta) obtinguts amb l'ajust al canal $B^+ \to J/\psi K^+$ amb 1 fb^{-1}.

Combinació dels algoritmes d'etiquetatge

Per tal de millorar l'eficiència efectiva en les mesures de violació de CP i de les oscil·lacions de mesons B, es poden combinar les decisions individuals dels diferents algoritmes en una sola decisió amb una major probabilitat d'encertar el sabor del quark. Per fer-ho utilitzem la probabilitat del *tagger* d'encertar el sabor del quark, definida com $p_i = 1-\omega_i$, i la decisió de cadascun d'ells (d_i), i les combinem de la forma següent:

$$\mathcal{P}(b) = \frac{p(b)}{p(b)+p(\bar{b})}, \qquad \mathcal{P}(\bar{b}) = 1-\mathcal{P}(b), \tag{8}$$

a on

$$p(b) = \prod_i \left(\frac{1+d_i}{2} - d_i p_i\right), \qquad p(\bar{b}) = \prod_i \left(\frac{1-d_i}{2} + d_i p_i\right). \tag{9}$$

D'aquesta forma, la decisió final es pren comparant els valors $\mathcal{P}(b)$ i $\mathcal{P}(\bar{b})$. Es poden realitzar també dues combinacions, només tenint en compte els OS *taggers* o bé incloent-hi també els SS. A causa de correlacions entre els OS *taggers*, és necessaria una correcció d'η. Això es fa mitjançant una calibració a posteriori de la combinació dels OS *taggers* amb dades segons l'Eq. 6. Un cop realitzada aquesta calibració es pot observar com els paràmetres p_0 i p_1 són compatibles amb $< \eta_c >$ i 1 respectivament, cosa que indica que la combinació ha estat calibrada correctament tal i com s'observa a la Taula 1.

channel	p_0	p_1	$< \eta_c >$
$B^+ \to J/\psi K^+$	0.392 ± 0.002	1.035 ± 0.021	0.391
$B^0 \to J/\psi K^*$	0.400 ± 0.004	1.013 ± 0.053	0.390
$B^0 \to D^-\pi^+$	0.398 ± 0.003	1.010 ± 0.039	0.393

TABLE 1: Valors de l'ajust de ω mesurats en els canals $B^+ \to J/\psi K^+$, $B^0 \to J/\psi K^*$ i $B^0 \to D^-\pi^+$ per la combinació dels OS *taggers* amb 1 fb^{-1}.

En la Figura 7 es pot veure, pel $B^0 \to D^{*-}\mu^+\nu_\mu$ com la distribució de l'ω estimada o η correspon exactament amb la calculada (ω).

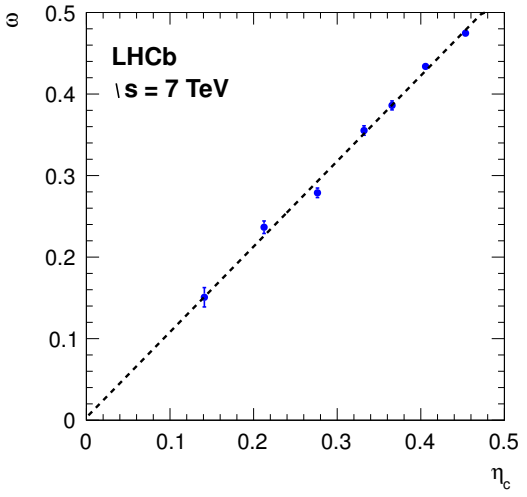


FIGURE 7: Distribució de ω versus η pel canal $B^0 \to D^{*-}\mu^+\nu_\mu$.

Un cop calibrada la combinació de l'OS, aquesta es pot combinar amb la decisió del SS mitjançant l'Eq. 8, ja que la correlació entre OS i SS és pràcticament nul·la.

S'han dissenyat altres alternatives per tal de corregir les correlacions entre els OS taggers. Però tot i estar a l'abast en els paquets FT i FTC, finalment no s'han utilitzat. Aquestes alternatives consistien en excloure el *tagger* que calcula la càrrega del vèrtex secundari quan hi ha un altre OS *tagger* present en l'esdeveniment, o bé, en la utilització d'una xarxa neuronal que té en compte les correlacions entre els *taggers*. Les dues alternatives estan explicades amb detall en les Seccions 5.3.1 i 5.3.2 respecivament.

Categories d'etiquetatge

Un cop es combinen els algoritmes i s'obté una decisió final i una ω, aquestes es poden utilitzar per calcular l'eficiència efectiva de la mostra considerant tots els esdeveniments per igual ("average") o bé categoritzant els esdeveniments en funció de la probabilitat que la decisió sigui correcta. En aquest últim cas, es poden classificar els esdeveniments en mostres independents en funció de la seva ω i obtenir una millora en l'eficiència efectiva de l'etiquetatge ("combine").

Per fer aquesta categorització s'utilitza la ω predita mitjançant la combinació lineal dels taggers de l'Eq. 8. També es va elaborar una estratègia amb MC que donava resultats similars. Aquesta classificava els esdeveniments segons els *taggers* actius i està detallada en la Secció 5.4.1.

Esdeveniment a esdeveniment

Havent aconseguit una ω de la combinació dels *taggers* calibrada amb dades reals, aquesta probabilitat d'error estimada per a cada esdeveniment es pot fer servir directament, maximitzant així l'eficiència efectiva. En aquest cas, la dilució ve donada per: $\sum_i (1-2\omega_i^2)/N$. A la Taula 2 es pot veure l'eficiència efectiva en diferents canals de control mitjançant diferents metòdes, justificant l'ús de l'ω esdeveniment a esdeveniment.

Channel	$\epsilon_{eff}^{average}\%$	$\epsilon_{eff}^{combine}\%$	$\epsilon_{eff}^{evt-p-evt}\%$
$B^+ \to J/\psi K^+$	1.69 ± 0.1	2.07 ± 0.11	2.10 ± 0.08
$B^0 \to J/\psi K^*$	1.24 ± 0.20	1.57 ± 0.22	2.09 ± 0.09
$B^0 \to D^{*-}\mu^+\nu_\mu$	1.58 ± 0.06	2.05 ± 0.06	2.53 ± 0.10

TABLE 2: Comparació de les eficiències efectives obtingudes en diferents canals de control amb 370 pb^{-1} fent servir l'ω "average", la "combine" i l'"event-per-event".

A més, aquesta ω és la que s'utilitza en la mesura de violació de CP en diferents decaïments maximitzant així l'ϵ_{eff} de l'etiquetatge de sabor.

Errors sistemàtics

Per tal de no comprometre la precisió de les mesures de violació de CP, s'ha de tenir un control dels errors sistemàtics. Aquests errors poden ser deguts bàsicament a les causes següents: asimetries en la producció d'hadrons amb quarks b i $\bar{b}$; asimetries entre partícules i antipartícules, a causa d'ineficiències del detector o a dependències a causa de la càrrega; així com també a un error en ω, o bé; a causa del soroll present en la selecció.

No es van detectar grans diferències en els resultats de l'etiquetatge en funció del sabor dels mesons, la càrrega de les partícules, la decisió de l'etiquetatge i en funció del camp magnètic del detector (Taula 6.1). També es van comparar els resultats de l'etiquetatge depenent de la multiplicitat de l'esdeveniment. Es va arribar a la conclusió que l'eficiència efectiva empitjora a mesura que la multiplicitat o el nombre de vèrtex primaris augmentaven (Taula 6.2).

Finalment es van calcular els errors sistemàtics en la calibració d'η a causa d'un canvi de polaritat de l'imant, en funció de la decisió d'etiquetatge i del sabor del mesó B, així com també en funció del model usat per descriure la distribució d'η i del nombre de PV. Aquests errors s'han d'introduir en els paràmetres de la calibració usats en altres estudis. Per l'OS i 1 fb^{-1} són:

$$p_0 = 0.392 \pm 0.002(\text{stat}) \pm 0.009(\text{syst}) \qquad \text{p}_1 = 1.035 \pm 0.021(\text{stat}) \pm 0.012(\text{syst}) \tag{10}$$

Resultats i conclusions

Gràcies al disseny dels algoritmes d'etiquetatge en MC i a la posterior optimització i calibració en dades, podem obtenir una informació versemblant del sabor del mesó B, així com, fer les primeres mesures de l'ϵ_{eff} o del poder de l'etiquetatge a LHCb garantint-ne el bon funcionament i la universalitat en diferents canals de control.

Els algoritmes d'etiquetatge de sabor avaluats als canals de control donen els resultats presentats a la Taula 3, que corresponen a la lluminositat de 1 fb^{-1} i als esdeveniments etiquetats per l'OS.

Channel	ϵ_{tag}%	ω %	ϵ_{tag} (1 - 2ω) %
$B^+ \to J/\psi K^+$	33.17 ± 0.09	36.7 ± 0.2	2.35 ± 0.06
$B^0 \to J/\psi K^*$	33.27 ± 0.14	36.7 ± 0.2	2.37 ± 0.06
$B^0 \to D^-\pi^+$	38.51 ± 0.15	38.29 ± 0.04	3.16 ± 0.24

TABLE 3: Resultats de l'etiquetatge de sabor pels canals $B^+ \to J/\psi K^+$, $B^0 \to J/\psi K^*$ i $B^0 \to D^-\pi^+$ mitjançant l'OS i 1 fb^{-1}.

En el cas del SS *taggers*, el SSK ha estat avaluat en el canal $B_s^0 \to D_s^-\pi^+$donant una $\epsilon_{\mathit{eff}} = 1.4 \pm 0.4\%$, mentre que pel SS$\pi$, on l'estratègia de calibració encara ha de ser definida, els resultats preliminars indiquen una ϵ_{eff} de $0.75 \pm 0.08\%$ i $1.12 \pm 0.08\%$ en els canals $B^0 \to J/\psi K^*$ i $B^0 \to D^-\pi^+$ respectivament.

A més, l'etiquetatge de sabor dels mesons B ha permès fer diferents anàlisis en el camp de la violació de CP, així com fer mesures d'oscil·lacions de mesons B. Algunes d'aquestes mesures han estat les més precises del moment, les més rellevants es troben resumides a l'appèndix B. Gràcies al bon funcionament de l'etiquetatge i a les dades recollides a finals del 2011, s'esperen nous resultats que poden descartar o confirmar indicis de nova física.

A les fades

Introduction

The Standard Model (SM) describes successfully the interactions of fundamental particles up to the actual energies, even though there is a number of evidences suggesting that it is only a special case of a more general theory. In this context, CP violation related effects can be very interesting for the discovery of new physics. To make precise CP violation measurements it is necessary to have a very accurate estimation of the flavour of the reconstructed B meson, a process known as *Flavour Tagging*.

The aim of this thesis is to provide a complete and detailed documentation of the Flavour Tagging at LHCb where it constitutes a basic tool for the measurement of B oscillations and for the study of CP violation. Its developement has been possible thanks to the effort of many people in the LHCb Collaboration, who have been working for many years. This, has allowed us to perform recently some of the best and most precise measurements ever made of the B oscillations as well as different and very precise analyses in the B sector of CP violation.

As member of the LHCb Collaboration my personal contribution has been the design of the flavour tagging algorithms (opposite side and same side) in MC, looking for new variables which would optimize the tagging performances, building the neural nets used to obtain the mistag fraction with the last MC simulations and also in the search for new taggers, like the D^0 tagger. In particular, I have also taken active part in the development of the vertex charge tagger and the optimization strategies of the taggers, first with the $B^0 \to D^{*-}\mu^+\nu_\mu$ channel by a fast fit to the flavour oscillation, then with the $B^+ \to J/\psi K^+$ channel, where the probability of mistag can be computed by comparing the tagging decision with the true flavour of the reconstructed B. I also developed an optimization method to be used with $B^0 \to D^-\pi^+$ events, and I participated to the calibration with $B^+ \to J/\psi K^+$ events. Finally, I have designed three additional methods which allowed us to combine the tagging algorithms. It is worth to mention that I have been the responsible for the Flavour Tagging package in the LHCb Collaboration.

This document is structured as follows: in Chapter 1 there is a short explanation of the theoretical motivations, describing the LHCb physics and what CP violation is. This is followed by Chapter 2 with a description of the LHCb detector that allowed us to make the measurements.

Chapter 3 describes in greater detail how the flavour tagging algorithms were designed on Monte Carlo events (for both same side and opposite side tagging algorithms). Since we do not want to rely on MC simulations alone, the tagging algorithms must be optimized and calibrated in real data to obtain the mistag fraction to be used in the CP fits. The different strategies that were used for this purpose are described in Chapter 4.

After their optimization and calibration, the individual tagging methods (known as *taggers*) are combined into a single decision. In Chapter 5, the default combination is explained in detail as well as other different combination strategies tested on MC. Additionally, in the same Chapter, the calibration of the OS taggers combination is described. This is used to correct biases in the wrong tag fractions due to correlations among OS taggers. Once the OS or OS + SS tagging decisions are combined, the systematic errors are evaluated and shown in Chapter 6.

Finally, last Chapter summarizes the results obtained with different control channels during the different physics runs before 2012. The readiness and good response of the tagging algorithms at the LHCb experiment allowed us to perform different measurements of B oscillations and in the CP field, some of them with the best precision ever obtained. These first measurements are summarized in the Appendices together with the description of the control channels used to obtain the tagging performances.

Chapter 1

LHCb Physics. CP Violation

The Standard Model (SM), introduced in 1961 by Glashow, Weinberg and Salam [1–3], gives a satisfying description of the electroweak interactions up to the actual energies (less than 1 TeV) but has some weak spots: it does not include gravity, it does not explain the mass hierarchy, the neutrino masses are zero, it introduces some free parameters, etc. Some aspects seem to indicate that the SM is the low energy realisation of a more general theory. New efforts are now being made to check the SM at a high precision and the search for proves of new physics beyond the SM has started. In this environment, CP violation plays a very important role: cosmological observations show an indirect excess of CP violation with respect to the SM predictions [4]. The CP violation related effects could be a very interesting point for the discovery of new physics.

This chapter deals with the following aspects: a brief introduction to the Standard Model and the CP violation, a discussion about the CKM mixing matrix, a development of the formalism for the B meson system and an explanation of the three types of CP violation.

1.1 The Standard Model

The Standard Model (SM) of particle physics is the best model to describe the interactions of fundamental particles. As the SM states, there are two types of fundamental particles: fermions and bosons, which are considered as matter particles and force carriers respectively.

The fermions are spin half integer particles and hence obey Fermi-Dirac statistics. They can be subdivided into quarks and leptons. Quarks are confined within bound states known as hadrons, which can be mesons or baryons. We call them mesons when they are formed by a quark-antiquark pair ($q\bar{q}$), but if they consist of bound states of three quarks or antiquarks (qqq or $\bar{q}\bar{q}\bar{q}$) they are called baryons. The leptons, which have intrinsic properties like electric charge, spin or mass, can be charged or neutral objects. Charged leptons can combine with

other particles to form various composite particles such as atoms, while the uncharged leptons, called neutrinos, rarely interact with anything and were postulated by Wolfgang Pauli in 1930 in order to preserve the conservation of energy, momentum and angular momentum in β decays.

There are three generations of quarks and leptons. The fact that there are only three generations is an experimental result, the first evidence comes from the ALEPH detector in the Large Electron Positron (LEP) accelerator [5]. This results in 6 discrete quark flavours (named: up, down, charm, strange, top -or truth- and bottom -or beauty-) with the corresponding antiquarks possessing the appropriate antiflavour, 3 charged leptons (electron, muon, tau) and one uncharged lepton per generation (named electronic, muonic and tauonic neutrino respectively). The fermions and their masses are listed below. Masses are taken from [6].

$$Leptons: \begin{pmatrix} e^- & 511.0 & KeV \\ \nu_e & < 2eV & \end{pmatrix} \begin{pmatrix} \mu^- & 105.7 & MeV \\ \nu_\mu & < 2eV & \end{pmatrix} \begin{pmatrix} \tau^- & 1776.8 & MeV \\ \nu_\tau & < 2eV & \end{pmatrix} \tag{1.1}$$

$$Quarks: \begin{pmatrix} \mathbf{u} & 1.5-3.3MeV \\ \mathbf{d} & 3.5-6.0MeV \end{pmatrix} \begin{pmatrix} \mathbf{c} & 1.27^{+0.07}_{-0.11}GeV \\ \mathbf{s} & 104^{+26}_{-34} \end{pmatrix} \begin{pmatrix} \mathbf{t} & 171.2 \pm 2.1GeV \\ \mathbf{b} & 4.20^{+0.17}_{-0.07}GeV \end{pmatrix} \tag{1.2}$$

On the other hand, bosons are particles with integer spin and obey the Bose-Einstein statistics. As mentioned before, the bosons are the carriers of the interactions, and will be explained in the next sections.

1.1.1 Strong and Electroweak Interactions

The matter particles, or fermions, interact via three forces in the SM through the bosons. The 3 forces are: the strong force, the weak force and the electromagnetic force; gravity is omitted in the SM due to its weakness. Interactions are introduced by requiring that the SM lagrangian exhibits local gauge invariance. Specifically, the theory is invariant under local gauge transformations of the group $SU(3) \times SU(2) \times U(1)$. The $SU(3)$ symmetry results of the interactions via the strong force, while the $SU(2) \times U(1)$ symmetry give rise to the electroweak interaction.

The strong and electroweak interactions are mediated by different type of bosons. The first one is mediated by gluons while the electroweak force is carried by the photon, the $W^\pm$ and the Z. Masses are given to these particles via their interaction with the Higgs field. Both interactions are described below.

Strong Interaction

The Strong Interaction is the force which binds the quarks within hadrons (the leptons do not interact via the strong force). The underlying theory behind the strong interaction is Quantum Chromodynamics (QCD), which acts only on objects with color, an additional quantum number assigned to each quark. Quarks can take three colors: red, green and blue (rgb) or the corresponding anticolors for antiquarks ($\bar{r}\bar{g}\bar{b}$), with the property that a bound state of all three color states (baryons), or any color and its anticolor (mesons), has no overall color, being mesons and baryons color neutral objects. The color structure of the SM can be described by the $SU(3)$ gauge group. Properties of this group give rise to 8 independent gauge bosons, known as gluons, which exist as color-anticolor states and can self-interact. These gluons mediate the strong interaction.

To explain an important particularity of the underlying theory behind the strong interaction (QCD), we need to first introduce the beta function $\beta(g_s)$ of quantum field theory that encodes the running of a coupling parameter, and it is defined by the relation:

$$\beta(g_s) = \mu \frac{\partial g_s}{\partial \mu} = \frac{\partial g_s}{\partial \ln \mu} \tag{1.3}$$

where μ is the energy scale of the given physical process and g_s is the runing coupling constant. If the beta functions of a quantum field theory vanish, then the theory is scale-invariant. However, in non-Abelian gauge theories, like QCD, this beta function can be negative, as first found by Frank Wilczek, David Gross and David Politzer [7, 8]. As a result, the QCD coupling decreases at high energies. Furthermore, the coupling decreases logarithmically, producing a phenomenon known as asymptotic freedom [9, 10], the discovery of which was awarded with the Nobel Prize in Physics in 2004 [11]. Conversely, the coupling increases with decreasing energy. This means that the coupling becomes large at low energies, and one can no longer rely on perturbation theory. By now, all these is very well established in experimental results. In the Figure 1.1 the effective coupling constant, called α_s, is shown as a function of energy.

In QCD asymptotic freedom means that the quarks behave as free particles (at least in terms of strong interaction) at short distances, but the strength of the strong interaction increases at larger distance, resulting in the confinement of quarks within hadronic colorless bound states.

Electroweak Interaction

Particles with non-zero electric charge interact via the Electromagnetic force, which is easily visible at a macrosopic level. The photon, without electric charge, is the associated gauge boson and is represented as γ. The theory describing the electromagnetic interactions is called

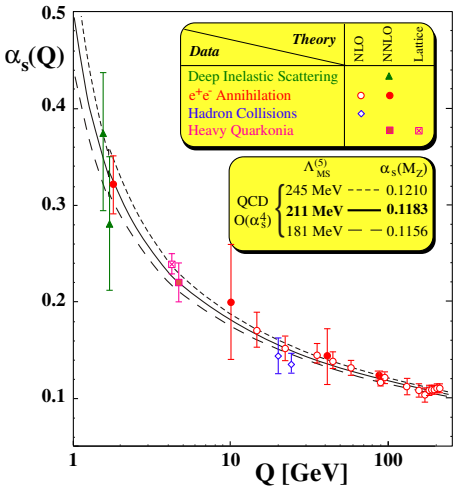


FIGURE 1.1: The QCD effective coupling constant as a function of the energy. A collection of different measurements by S. Bethke [12] is shown.

Quantum Electrodynamics (QED) and is associated to the gauge symmetry generated by the $U(1)$ group.

The weak interaction operates only on left-handed particles (the handedness of a particle refers to the projection of the spin of a particle relative to the direction of its momentum), this is a consequence of the $V-A$ (Vector minus Axial) form of the interaction as the lagrangian contains terms which projects out the left-handed component of the state [13].

The electromagnetic and weak interactions are unified into the $SU(2)_L \times U(1)_Y$ gauge symetry and the associated massless gauge bosons are $\vec{W}_\mu = (W^1_\mu, W^2_\mu, W^3_\mu)$ and B_μ. The Higgs mechanism breaks this symmetry and decouples the weak and the electromagnetic interactions giving rise to the photon (gauge field A_μ) and the three massive weak gauge bosons:

$$A_\mu = -\sin\theta_W W^3_\mu + \cos\theta_W B_\mu \tag{1.4}$$

$$W^\pm_\mu = \frac{W^1_\mu \pm i W^2_\mu}{\sqrt{2}} \tag{1.5}$$

$$Z_\mu = \cos\theta_W W^3_\mu + \sin\theta_W B_\mu \tag{1.6}$$

where θ_W is the Weinberg angle, defined as: $\tan\theta_W = \frac{g'}{g}$, where g' and g are the $U(1)_Y$ and $SU(2)_L$ couplings. The weak interaction term that couples the charged gauge bosons to the fermions is:

$$\mathcal{L}_{int} = -\frac{g}{\sqrt{2}} W^+_\mu (\bar{\nu}\gamma^\mu(1-\gamma^5)l + \bar{q}_u\gamma^\mu(1-\gamma^5)q_d + h.c). \tag{1.7}$$

ν, l, q_u and q_d are the column vectors for $(\nu_e, \nu_\mu, \nu_\tau)$, (e^-, μ^-, τ^-), (u, c, t) and (d, s, b) respectively.

In order to accomodate the massive gauge bosons, the Higgs mechanism breaks the symmetry. The introduction of a scalar field, with a potential $V(\Phi)$, chosen to be equal to $-\mu^2|\Phi|^2+\lambda^2|\Phi|^4$, keeps the renormalization of the theory because while the lagrangian remains symmetric, the symmetry of the theory is broken by choosing spontaneously one of the degenerate ground states as the true ground, resulting in the appearance of massless Goldstone bosons. For practical reasons the minimum of the potential is chosen as:

$$\Phi(x) = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ \frac{2\mu}{\sqrt{\lambda}} + h(x) \end{pmatrix} \tag{1.8}$$

With the Yukawa term of the Standard Model lagrangian, which couples the Higgs boson to the fermions, the Higgs mechanism gives the Goldstone bosons mass. This lagrangian mass term has this form:

$$\mathcal{L}_{mass} = -\frac{v}{\sqrt{2}}\left(\bar{l}M_l l + \bar{q}_u M_u q_u + \bar{q}_d M_d q_d\right) \tag{1.9}$$

where M_l, M_u and M_d are the mass matrices (in the Standard Model neutrinos are massless). If these matrices are diagonalized, to put the Standard Model lagrangian in terms of the mass eigenstates, it results that these are not the same as the flavor eigenstates as in equation 1.7. A unitary matrix should be introduced in order to rewrite the lagrangian in function of the quark mass eigenstates. This matrix is the CKM matrix that gives rise to CP violation [14].

1.2 CP Violation

The CP transformation combines charge (C) conjugation with parity (P). Under C, particles and antiparticles are interchanged. Under P, the handedness of the space is reversed, $\overrightarrow{x} \rightarrow -\overrightarrow{x}$. It was once believed that physics was invariant under these discrete symmetries, but in 1956, C.S Wu observed that the mirror image of the left-handed neutrino, the right-handed neutrino, does not exist and that therefore the symmetry of the weak interaction is broken by the P operation [15]. Charge conjugation, C, replaces a left-handed neutrino with a left-handed anti-neutrino. It can be seen too that C is not conserved. The symmetry is restored when the P operator is not applied alone, but when the combined operation CP is applied. CP transformation transforms a left-handed neutrino into a right-handed anti-neutrino, which does exist.

In 1964, a symmetry violation of the CP transformation was observed by James Cronin and Val Fitch in the case of the neutral K-meson at a level of 0.2% [16]. It was observed in neutral B

decays in 2001 at the BaBar [17] and Belle [18] collaborations. A first indication of CP violation in neutral D mesons decays has been recently reported by the LHCb Collaboration [19]. The present measurements of CP asymmetries provide some of the strongest constrains in the weak couplings of quarks. The expectation is that future measurements of CP violation will provide additional constrains on the flavour parameters of the SM and can prove new physics beyond the SM.

1.2.1 CP violation in quantum field theory

In quantum field theory (QFT) scalar fields $\phi(x)$, vector fields $V(x)$, and fermion fields $\psi(x)$ can be transformed under C, P and T as shown in [20, 21]. When a CP operator is applied, every field is converted into its adjoint partner, and only if the operator completely converts the two adjoint lagrangian terms into each other, the CP symmetry is satisfied.

If the lagrangian contains complex constants (as from hermiticity the coupling constant of one adjoint term is the complex conjugate of the matching order) CP violation can be easily accomodated. However, CP symmetry can still be conserved when the complex constants are absorbed through redefinitions of the quantum fields, as the physical observables are invariant under phase transformations, i.e. $\psi(x) \to e^{i\phi}\psi(x)$. After a number of phase redefinitions, should one or more phases be irreductible and cannot be absorved, the theory is not invariant under CP.

In the SM the kinetic gauge term and the Higgs term that gives rise to spontaneous symmetry breaking of the lagrangian are automatically CP-invariant and it is therefore the Yukawa sector (which describes the interaction between the fermionic and scalar fields) which must be the source of CP violation, introduced through complex Yukawa couplings.

The Yukawa couplings form a 3×3 mass matrix for the up and down type quarks, $u^{(weak)}$ and $d^{(weak)}$ respectively, which are eigenstates of the weak interaction. These matrices can be defined in the basis of mass eigenstates as $u_i^{(weak)} = U_{ij}^{(u)} u_j^{(mass)}$ and $d_i^{(weak)} = U_{ij}^{(d)} d_j^{(mass)}$, where $U^{(u,d)}$ are rotation matrices. Neutral weak interactions are unaffected ($\bar{u}_i^{(weak)} u_i^{(weak)} = \bar{u}_i^{(mass)} u_i^{(mass)}$) as the transformation is unitary, while the charged weak interactions transforms as:

$$\bar{u}_i^{(weak)} d_i^{(weak)} \to \bar{u}_i^{(mass)} (U^{(u)})^\dagger U^d d_i^{(mass)} \tag{1.10}$$

The elements of the unitary matrix $\mathbf{V_{CKM}} \equiv (U^{(u)})^\dagger U^d$, are proportional to the strength of the couplings between the up and down type quarks. It is known as the Cabibbo-Kobayashi-Maskawa (CKM) matrix [22] and it comprises a complex phase which is the source of CP violation in the SM.

1.3 The CKM Matrix

The quarks that interact via the weak interaction are superpositions of mass eigenstates. The Cabibbo-Kobayashi-Maskawa (CKM) matrix links the mass eigenstates for the down-type quarks (d,s,b), with the weak eigenstates ($\tilde{d}$,$\tilde{s}$,$\tilde{b}$). It is a complex matrix with a priori $2n^2$ real parameters (n as the number of quark families), but from unitarity one can see that only $(n-1)^2$ parameters are independent. In the SM $n = 3$, the CKM matrix can be described by four real parameters; these corresponds to three rotation angles θ_{ij} and a complex phase δ, which is the source of CP violation in the SM.

$$
\begin{aligned}
V_{CKM} &\equiv \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \\
&= \begin{pmatrix} c_{12}c_{23} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}
\end{aligned} \tag{1.11}
$$

where $s_{ij} = \sin\ \theta_{ij}$ and $c_{ij} = \cos\ \theta_{ij}$. Using the sine of the Cabibbo angle ($\lambda = \sin\ \theta_{12} \simeq 0.22$) as an expansion parameter, the CKM matrix can be writen in the Wolfenstein parametrization [23] as follows:

$$
V_{CKM} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \tag{1.12}
$$

where A, ρ and η are the remaining three parameters. In this convention CP violation is present if $\eta \neq 0$. The unitarity of the CKM matrix ($V^\dagger V = V V^\dagger = \mathbb{I}$) implies various relations among its elements. There are 6 orthogonality conditions:

$$V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0, \tag{1.13}$$

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0, \tag{1.14}$$

$$V_{ud}V_{us}^* + V_{cd}V_{cs}^* + V_{td}V_{ts}^* = 0, \tag{1.15}$$

$$V_{ud}V_{td}^* + V_{us}V_{ts}^* + V_{ub}V_{tb}^* = 0, \tag{1.16}$$

$$V_{cd}V_{td}^* + V_{cs}V_{ts}^* + V_{cb}V_{tb}^* = 0, \tag{1.17}$$

$$V_{ud}V_{cd}^* + V_{us}V_{cs}^* + V_{ub}V_{cb}^* = 0. \tag{1.18}$$

Each of these relations requires the sum of the three complex quantities to vanish and so they can be geometrically represented in the complex plane as triangles. They are known as the

'Unitarity Triangles'. All of the triangles have the same area [24] and only two have sides of roughly equal lengths (Eq. 1.14 and 1.16). These two triangles (Figure 1.2) are especially important in the study of CP, the relative length of their sides means that measurements of the internal angles are feasible, providing tests of CKM physics, that can lead to potentially large CP violating asymmetries from phases between CKM matrix elements. The other triangles have at least one side much shorter than the other two, and so they would almost collapse in a line. This gives an intuitive understanding of why CP violation is small in the leading K (1.15) and B_s (1.13) decays.

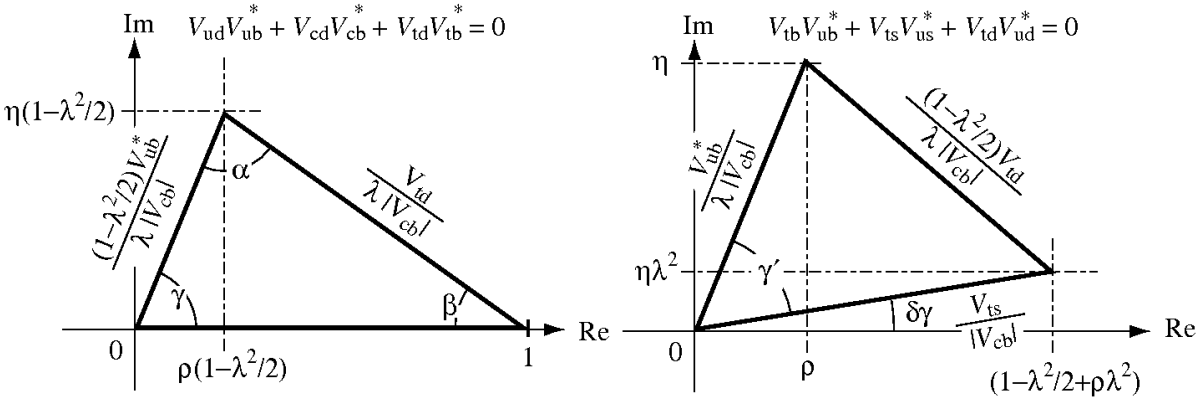


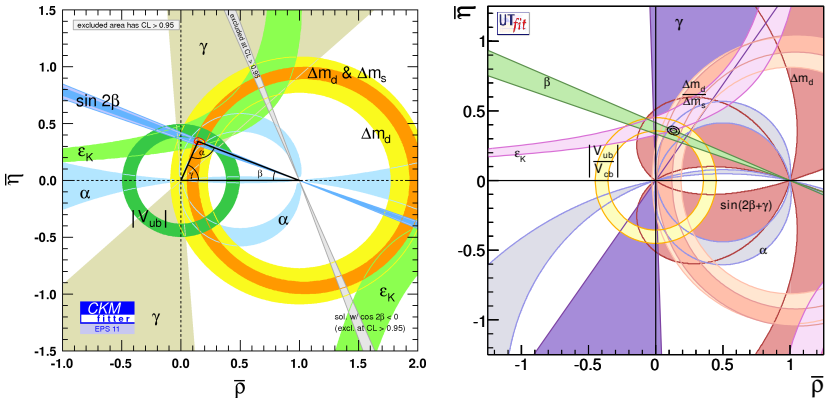
FIGURE 1.2: The unitarity triangles come from equations 1.14 (left) and 1.16 (right). Both triangles have been scaled by a factor $V_{cd}V_{cb}^*$ so one side of the Unitarity Triangle lies along the real axis.

The triangle represented in Figure 1.2 (left) is known as the 'Unitarity Triangle' and represent the most experimentaly accessible of the unitarity relations, which involves the two smallest elements of the CKM matrix. The three angles of the Unitarity Triangle, α, β and γ are given by:

$$\alpha = \arg\left[\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right], \quad \beta = \arg\left[\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right], \quad \gamma = \arg\left[\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right], \tag{1.19}$$

where $\alpha + \beta + \gamma = \pi$. These three angles offer a test for the SM flavour structure, thus an unclosed Unitarity Triangle would be a sign of new physics beyond the SM. Constrains on these angles can be extracted from measurements of many processes, and combined fits based on current best measurements of SM parameters have been performed by both the CKMfitter [25] and UTfit [26] groups. The combined fit results before any LHCb measurement can be seen in Figure 1.3, and the best fit value for α, β and γ were [27]:

$$\alpha = (91.0 \pm 3.9)^\circ \qquad \beta = (21.76^{+0.92}_{-0.82})^\circ \qquad \gamma = (67.2 \pm 3.9)^\circ \tag{1.20}$$



(a) CKM Fitter (b) UT Fit

FIGURE 1.3: Combined fit results of the Unitarity Triangle.

The CP violation effects in the B_s mesons are small. Its associated triangle (Figure 1.4) has less experimental constrains, but the CP violation can be determined through the measurement of the β_s angle, which is defined as:

$$\beta_s \equiv \left(-\frac{V_{cb}V_{cs}^*}{V_{tb}V_{ts}^*}\right) \tag{1.21}$$

which is equivalent to the β angle, but for B_s^0 mesons. The actual indirect determination via global fit to experimental data ($\beta_s = 0.0182 \pm 0.0009$ [25]) is compatible with the SM predictions. This parameter is accesible for example via $B_s^0 \to J/\psi\phi$ decays from which one has:

$$\phi_s^{J/\psi\phi} = -2\beta_s + \delta^{penguin} + \phi^{NP} \tag{1.22}$$

where $\delta^{penguin}$ refers to penguin contributions, which are though to be negligible, and ϕ^{NP} is an eventual phase from new physics. The actual best measurement of the $\phi_s^{J/\psi\phi}$ has been performed by the LHCb collaboration obtaining $\phi_s^{J/\psi\phi} = 0.15 \pm 0.18 \pm 0.07$ at a 68% confidence level compatible with the SM at 1σ [28].

In Table 1.1 the values before any LHC measurement of CKM elements are taken from the Particle Data Group (PDG) [6] together with the procedure of the measurement.

1.4 The B meson system

The understanding of the B meson system is a key point for the measurement of the CKM parameters. Decays of B mesons, which contains a b quark or a $\bar{b}$ antiquark, can provide theoretically clean information about the angles of the Unitarity Triangle.

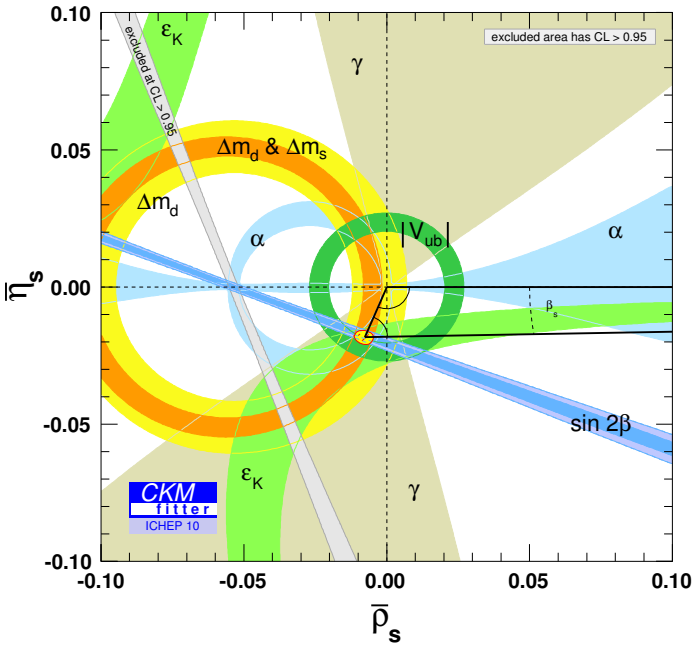


FIGURE 1.4: The unitarity triangle from equation 1.13 which concerns the B_s mesons.

Element	Value	Measurement Channel
$\|V_{ud}\|$	0.97425 ± 0.00018	Nuclear beta decays
$\|V_{us}\|$	0.22543 ± 0.00077	Semileptonic kaon decays
$\|V_{cd}\|$	0.22529 ± 0.00077	Neutrino scattering from valence d quarks
$\|V_{cs}\|$	$0.97342\ ^{+0.00021}_{-0.00019}$	Semileptonic D meson decays
$\|V_{cb}\|$	$0.04128\ ^{+0.00058}_{-0.00129}$	Semileptonic B meson decays
$\|V_{ub}\|$	$0.00354\ ^{+0.00016}_{-0.00014}$	Semileptonic B meson decays
$\|V_{td}\|$	$0.00858\ ^{+0.00030}_{-0.00034}$	B^0 mixing assuming $\|V_{tb}\| = 1$
$\|V_{ts}\|$	$0.04054\ ^{+0.00057}_{-0.00129}$	B^0_s mixing assuming $\|V_{tb}\| = 1$
$\|V_{tb}\|$	$0.99914\ ^{+0.00005}_{-0.00003}$	Single-top-quark production

TABLE 1.1: The current experimental status of the CKM matrix elements [6].

The b quark was first discovered in 1977 in the decays of the bound state $\Upsilon(1S) \equiv (b\bar{b})$ [29, 30] and the average lifetime of hadrons containing a b quark was first measured in 1983 [31, 32].

Weak interactions permit a phenomenon called B oscillations, which allows a B^0 meson to become a $\bar{B}^0$ and viceversa. $B^0\bar{B}^0$ oscillation provides two interfering amplitudes, which can lead to CP violation under the right conditions. Here and in the following section will be discussed the structure of the B meson system and how it behaves under CP violation.

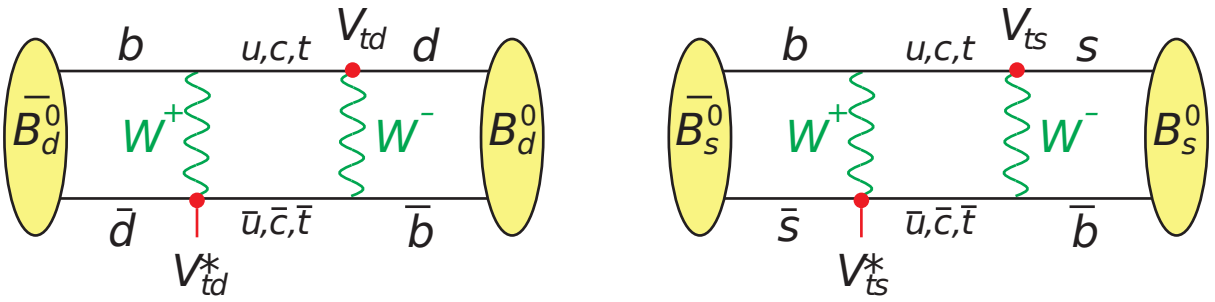


FIGURE 1.5: Box diagram for $B^0\bar{B}^0$ oscillation, for B_d and B_s mesons.

1.4.1 B^0 - $\bar{B}^0$ oscillation

The $B^0\bar{B}^0$ oscillations were first observed in 1987 [33, 34]. This process is allowed through box diagrams in the weak interactions as shown in Figure 1.5. However, the flavour eigenstates $|B^0\rangle$ and $|\bar{B}^0\rangle$, produced in quark-level strong interactions, are not eigenstates of the weak interactions. The eigenstate of the total hamiltonian, with definite mass and lifetime, are mixtures of the flavour eigenstates. The 'light' B_L and 'heavy' B_H mass eigenstates of the neutral B^0 meson system are given by:

$$|B_L\rangle = p|B^0\rangle + q|\bar{B}^0\rangle, \qquad |B_H\rangle = p|B^0\rangle - q|\bar{B}^0\rangle \tag{1.23}$$

where B^0 and $\bar{B}^0$ are the flavour eigenstates of the system, related through CP transformation according to: $\mathrm{CP}|B^0\rangle = e^{2i\beta}|\bar{B}^0\rangle$ (being β an arbitrary phase), and where the complex coefficients are normalized, i.e. $|p|^2 + |q|^2 = 1$. The phase of q/p, which depends on phase conventions, is not an observable, only the modulus $|q/p|$ has a physical significance.

In the Wigner-Weisskopf approximation [35], the time evolution and decay of the B^0 meson system is written in terms of only the flavour eigenstates, i.e. $|\psi(t)\rangle \equiv a(t)|B^0\rangle + b(t)|\bar{B}^0\rangle$. The wave function evolves according to the Schrödinger equation

$$i\frac{d}{dt}\begin{pmatrix} a \\ b \end{pmatrix} = \mathbf{H}\begin{pmatrix} a \\ b \end{pmatrix} = (\mathbf{M} - \frac{i}{2}\mathbf{\Gamma})\begin{pmatrix} a \\ b \end{pmatrix} \tag{1.24}$$

where the mass and lifeterm matrices $\mathbf{M}$ and $\mathbf{\Gamma}$ are Hermitian. The matrices can be computed from the weak hamiltonian (H_W) in second-order perturbation theory as

$$M_{ij} = m_B\delta_{ij} + \langle i|H_W^{\Delta B=2}|j\rangle + \Sigma_n \frac{1}{m_B - E_n}\langle i|H_W^{\Delta B=1}|n\rangle\langle n|H_W^{\Delta B=1}|j\rangle \tag{1.25}$$

$$\Gamma_{ij} = 2\pi\Sigma_n\delta(E_n - m_B)\langle i|H_W^{\Delta B=1}|n\rangle\langle n|H_W^{\Delta B=1}|j\rangle \tag{1.26}$$

Physical or on-shell states contribute to $\mathbf{\Gamma}$, whereas virtual or off-shell intermediate states contributes to $\mathbf{M}$. The B_L and B_H mass eigenstates have the complex eigenvalue $\mathbf{M} - i/2\mathbf{\Gamma} \pm \sqrt{H_{12}H_{21}}$ with $\mathbf{\Gamma} \equiv \Gamma_{11} \equiv \Gamma_{22}$ and $\mathbf{M} \equiv M_{11} \equiv M_{22}$. The CPT invariance guarantees $H_{11} = H_{22}$.

The mass and lifetime difference between the 'heavy' and 'light' eigenstates are

$$\Delta m_d - \frac{i}{2}\Delta\Gamma = (m_H - m_L) - \frac{i}{2}(\Gamma_H - \Gamma_L) = \sqrt{H_{12}H_{21}} \tag{1.27}$$

and the ratio q/p is expressed as

$$\frac{q}{p} = \sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}} \tag{1.28}$$

If we consider the mass difference ($\Delta m_d = m_{B_H} - m_{B_L}$) much bigger than the width difference between the two eigenstates ($\Delta\Gamma_d = \Gamma_{B_H} - \Gamma_{B_L}$), the two mass eigenstates have virtually equal lifetimes but noticeably different masses, and the equations 1.27 and 1.28 can be simplified to $\Delta m_d = 2|M_{12}|$ and $q/p = -|M_{12}|/M_{12}$. In terms of the CKM elements $q/p = e^{-2i\beta}$. In that case the time evolution of an initially pure B^0 or $\bar{B}^0$ state can be writen as follows:

$$|B^0(t)\rangle = e^{-imt}e^{\Gamma t/2}\{\cos(\Delta m_d t/2)|B^0\rangle + i(q/p)\sin(\Delta m_d t/2)|\bar{B}^0\rangle\} \tag{1.29}$$

$$|\bar{B}^0(t)\rangle = e^{-imt}e^{\Gamma t/2}\{\cos(\Delta m_d t/2)|\bar{B}^0\rangle + i(p/q)\sin(\Delta m_d t/2)|B^0\rangle\} \tag{1.30}$$

where $m = 1/2(m_{B_H} + m_{B_L})$ and $\Gamma = 1/2(\Gamma_{B_H} + \Gamma_{B_L})$. Equations 1.29 and 1.30 demonstrate that the probability of a B^0 to become a $\bar{B}^0$, or viceversa, oscillates as a funtion of time and depends on the mass difference Δm_d. The best value before any LHCb measurement for the mass difference was $\Delta m_d = 0.507 \pm 0.005$ ps^{-1} [6], and for the B_s meson $\Delta m_s = 17.77 \pm 0.10 \pm 0.07$ ps^{-1} [6]. LHCb has measured in 2011 Δm_d and Δm_s with an error similar to the previous measurements using hadronic decays $\Delta m_d = 0.499 \pm 0.032$ (stat) ± 0.003 (syst) ps^{-1} [36] and $\Delta m_s = 17.63 \pm 0.11$ (stat) ± 0.02 (syst) ps^{-1} [37]. As soon as the further data collected in the 2011 physics run is analysed, the LHCb will make world leading measurements on both Δm_d and Δm_s.

A useful parameter for characterising the mixing of neutral mesons is the ratio $x = \Delta m/\Gamma$, which is the magnitude of the oscillation frequency of the B meson mixing. It indicates how fast the oscillation between B^0 and $\bar{B}^0$ happens. For B_d mesons the ratio is small, $x_d = 0.774 \pm 0.008$, while in the B_s mesons, $x_s = 26.2 \pm 0.5$ [6]. The larger value (due to the larger mass difference) of x_s leads to a much more rapid oscillation between B_s^0 and $\bar{B}_s^0$. In Figure 1.6 it can be seen the effect of a $B^0\bar{B}^0$ osillation according to the B meson type.

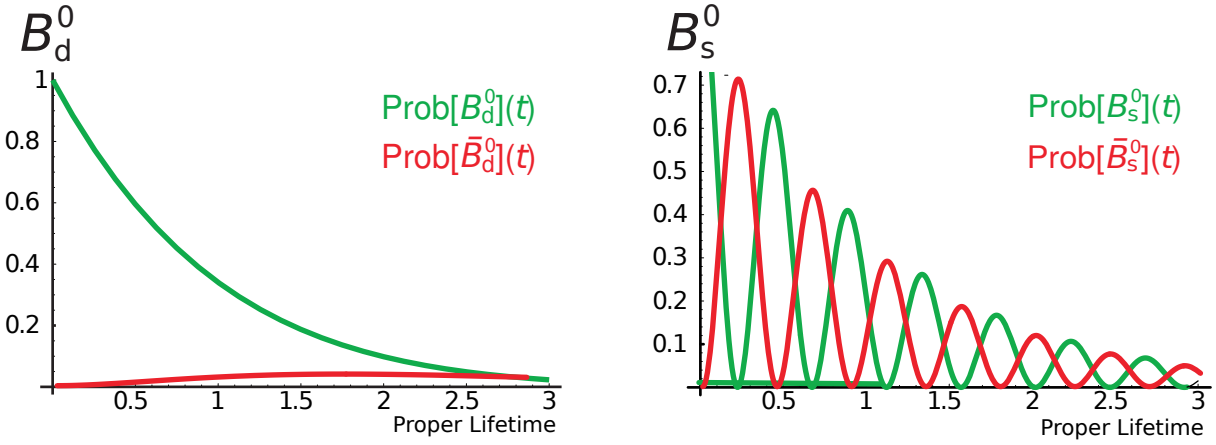


FIGURE 1.6: With an initial B^0_d (left) or a B^0_s (right), the green and red lines represent the probability of having a $B^0_{d,s}$ or a $\bar{B}^0_{d,s}$ as a function of the proper time.

1.5 Types of CP Violation

CP violation corresponds, as mentioned, to differences between processes and their CP conjugate processes. Since CP conjugate states are related via hermitian conjugates, the presence of complex couplings in the Hamiltonian will break invariance under CP transformation. The complex phase in the CKM matrix is an example of this. Physical observables are determined by the product of two quantum mechanical transition amplitudes, and thus, only phase differences have physical consequences. So CP violation can only be observed in processes that have multiple coherent distributions with different phases. The CP violation in meson decays can manifest itself in several ways, generally it is classified as one of three types:

- Direct CP violation, or CP violation in the Decay
- Indirect CP violation, or CP violation in Mixing
- CP violation in the Interference of Mixing and Decay

These categories are not mutually exclusive, and all can contribute to the overall CP violation observed in a single process.

1.5.1 Direct CP Violation

Direct CP violation (or CP violation in the decay) is due to the interference among decay amplitudes which differ in both weak and strong phases. Direct CP violation is the only type of CP violation for charged modes, while for neutral modes it competes with the other two types of CP violation.

For B decays, one builds time-independent CP asymmetry observables:

$$\mathcal{A}_{CP} = \frac{\Gamma(\overline{B} \to \overline{f}) - \Gamma(B \to f)}{\Gamma(\overline{B} \to \overline{f}) + \Gamma(B \to f)} = \frac{1 - |\overline{A}_{\overline{f}}/A_f|^2}{1 + |\overline{A}_{\overline{f}}/A_f|^2} \tag{1.31}$$

where $A_f = \langle f|H|B\rangle$ and $\overline{A}_{\overline{f}} = \langle \overline{f}|H|\overline{B}\rangle$ are the decay amplitudes. Sizeable direct CP violation effects ($|\overline{A}_{\overline{f}}/A_f| \neq 1$) require the contribution to the decay of at least two amplitudes of comparable size with of course different weak phases, but also a non-zero relative strong phase.

A direct CP violation has been observed in the decay of B mesons. $A_{CP}^{dir}(B_d^0 \to K^+\pi^-)$ has been measured at LHCb as $-0.074 \pm 0.033 \pm 0.008$ [38] which is compatible with the current world average -0.098 ± 0.013 [6].

1.5.2 Indirect CP Violation

Also known as CP violation in the mixing, it manifest itself in the neutral mesons mixing, as a consequence of the mass eigenstates being different from the CP eigenstates. Neutral mesons (K, D, B mesons) can oscillate between particle and antiparticle states via flavour-changing neutral processes, some of them are shown in Figure 1.7. The mixing phenomenon is closely related to CP violation, an example is the asymmetry in the transitions $B^0 \to \bar{B}^0$ and $\bar{B}^0 \to B^0$.

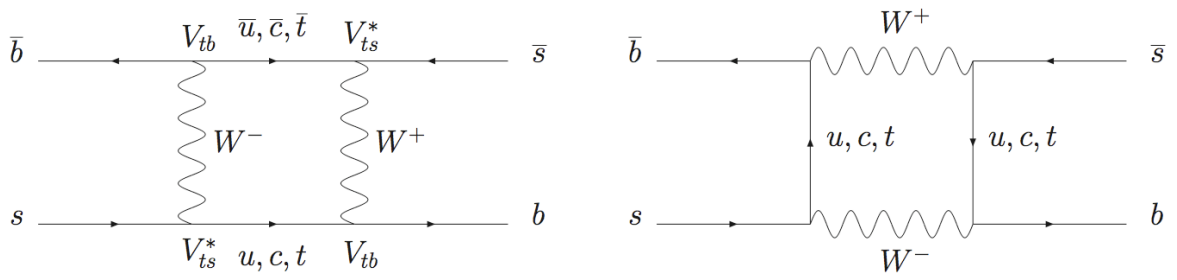


FIGURE 1.7: Two possible box diagrams for B^0 - $\bar{B}^0$ mixing.

The evolution of a physical B meson state can be described by the equations:

$$|B(t)\rangle = g_+(t)|B\rangle + \frac{q}{p} g_-(t)|\bar{B}\rangle \tag{1.32}$$

$$|\bar{B}(t)\rangle = g_+(t)|\bar{B}\rangle + \frac{p}{q} g_-(t)|B\rangle \tag{1.33}$$

Where parameters $g_\pm$ represent the time-dependent probabilities of the state remaining unchanged (+) or oscillating into its charge conjugate state (-), and are defined as:

$$|g_\pm(t)|^2 = \frac{e^{-\Gamma t}}{2}\left[\cosh\left(\frac{\Delta\Gamma}{2}t\right) \pm \cos(\Delta m\ t)\right] \tag{1.34}$$

where the Δm is the mass difference between the 'light' and 'heavy' mass eigenstates, and $\Delta\Gamma$ is the lifetime difference. The p and q coeficients denote the relative proportions of B and $\bar{B}$ states making up the mass eigenstates and play a fundamental role in Indirect CP violation. If $p = q = 1/\sqrt{2}$, the physical mass eigenstates are also CP eigenstates. If p and q are not equal, then CP violation can occur, therefore CP is violated if $|q/p| \neq 1$.

$$|q/p| \neq 1 \Longrightarrow Prob(B^0 \to \bar{B}^0) \neq Prob(\bar{B}^0 \to B^0) \tag{1.35}$$

This effect can be studied by investigating time-dependent differences in mixing rates in decays to flavour-specific final states such as semileptonic neutral B decays:

$$\mathcal{A}_T = \frac{\Gamma(|\bar{B}^0(t)\rangle \to l^+\nu X) - \Gamma(|B^0(t)\rangle \to l^-\overline{\nu} X)}{\Gamma(|\bar{B}^0(t)\rangle \to l^+\nu X) + \Gamma(|B^0(t)\rangle \to l^-\overline{\nu} X)} \tag{1.36}$$

As the proper time cancels out, the asymmetry is independent of t:

$$\mathcal{A}_T = \frac{1 - |q/p|^4}{1 + |q/p|^4} \tag{1.37}$$

The CP asymmetry in the mixing has been measured in semileptonic decays, and the current average [6] is: $\mathcal{A}_T^d = -0.0005 \pm 0.0056$ and $\mathcal{A}_T^s = -0.0049 \pm 0.0038$ for B_d^0 and B_s^0 mesons respectively.

Measuring CP violation in mixing can constrain new physics. New physics can potentially lead to greater CP violation in mixing than the standard model, caused by new particles that could enter inside the box diagram (Figure 1.7). Some models predicts bigger effects than the SM, up to two orders of magnitude [43].

1.5.3 CP Violation from Interference

CP violation can still occur in a process, even if there is no CP violation in mixing nor in decay individually, due to interference between their phases. To demonstrate this, we can calculate the time-dependent decay rates to a specific final state f accessible to both B^0 and $\bar{B}^0$ decays. For that, one introduces the phase-independent complex parameter λ_f, defined as [35]:

$$\lambda_f = \frac{q}{p}\frac{\overline{A}_f}{A_f} \tag{1.38}$$

The parameter λ_f is invariant under arbitrary rephasing of the initial and final states, and so it is a potential observable in neutral mesons decays. CP violation appears when $\lambda_f \neq \pm 1$. However, it is possible to satisfy $\lambda_f \neq \pm 1$ even if $|q/p| = 1$ and $|A_f/\bar{A}_{\bar{f}}| = 1$ if $Im(\lambda_f) \neq 0$.

$$\lambda_{f_{CP}} \neq \pm 1 \Longrightarrow Prob(|B^0(t)\rangle \to f_{CP}) \neq Prob(|\bar{B}^0(t)\rangle \to f_{CP}) \tag{1.39}$$

If we consider decays to CP eigenstates $|f\rangle$, such states are available in both B^0 and $\bar{B}^0$ decays. An example is the decay $B_d^0 \to J/\psi K_S^0$, which is depicted in Figure 1.8. The two paths interfere and produce a time dependent asymmetry in the decay rates $\Gamma_{B\to f(t)}$ and $\Gamma_{\bar{B}\to f(t)}$:

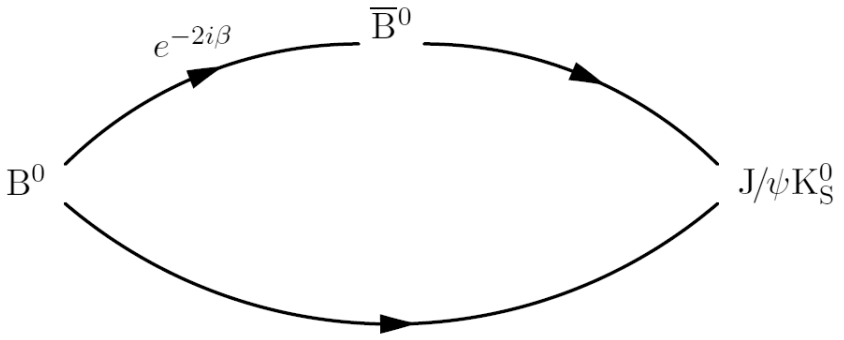


FIGURE 1.8: A schematic view of the $B_d^0 \to J/\psi K_S^0$ decay. The mixing introduces a phase difference between the two paths, which leads to CP violation.

$$\mathcal{A}_{f_{CP}}(t) = \frac{\Gamma(|\bar{B}^0(t)\rangle \to f_{CP}) - \Gamma(|B^0(t)\rangle \to f_{CP})}{\Gamma(|\bar{B}^0(t)\rangle \to f_{CP}) + \Gamma(|B^0(t)\rangle \to f_{CP})} \tag{1.40}$$

The asymmetry can be written as:

$$\mathcal{A}_{f_{CP}}(t) = S_{f_{CP}} \sin(\Delta m_d t) - C_{f_{CP}} \cos(\Delta m_d t) \tag{1.41}$$

where the coefficients of the sine and cosine terms are:

$$S_{f_{CP}} = \frac{2Im(\lambda_{f_{CP}})}{1 + |\lambda_{f_{CP}}|^2} \qquad \text{and} \qquad \mathrm{C_{f_{CP}}} = \frac{1 - |\lambda_{\mathrm{f_{CP}}}|^2}{1 + |\lambda_{\mathrm{f_{CP}}}|^2} \tag{1.42}$$

The cosine terms vanishes in absence of both CP violation in mixing ($|q/p| = 1$) and direct CP violation in the decay ($|\overline{A}_{f_{CP}}/A_{f_{CP}}| = 1$). Even in that case, CP violation can arise from the weak phase difference between q/p and $\overline{A}_{f_{CP}}/A_{f_{CP}}$, resulting in a non-vanishing sine term ($Im(\lambda_{f_{CP}}) \neq 0$):

$$S_{f_{CP}} = Im(\lambda_{f_{CP}}) = \eta_{f_{CP}} \sin 2\beta \tag{1.43}$$

with $\eta_{f_{CP}}$ being the CP parity of the f_{CP} final state (this last expression, for the SM case) and β the angle defined in Eq. 1.19. This is interpreted as an interference between the decay of a B^0 meson with and without mixing (i.e. $B^0 \to \bar{B}^0 \to f_{CP}$ and $B^0 \to f_{CP}$).

In the SM $C_{f_{CP}}$ is predicted to be zero for $b \to c\bar{c}s$ decays and it has been measured compatible with zero in the $B^0 \to J/\psi K_S$ decay by Belle [39] and BaBar [40] collaborations. The world average measurement before the LHCb measurement for $S_{J/\psi K^0_S} = \sin 2\beta$ was 0.673 ± 0.023 [41]. LHCb finds a compatible value ($\sin 2\beta = 0.53^{+0.28}_{-0.29} \pm 0.08$ [42]) dominated by the statistical uncertainty.

1.6 Implications beyond the SM

The SM summarizes successfuly most of our current understanding on particles physics. Even though, there is an increasing number of evidences (both experimental indications and theoretical considerations) suggesting that the SM is only a special case of some more general theory; the Hierarchy problem, the mechanism of baryogenesis, the lack of the gravity, the origin of neutrino masses, the strong CP problem, etc. These phenomena remain today still unexplained, that is why new models have been proposed to solve some of these problems.

- **Supersymmetry (SUSY)** is the most known model. It predicts the existence of a bosonic partner to every fermion and viceversa. The introduction of this symmetry causes the cancellation of the high-order corrections to the Higgs mass, allowing it to occupy the expected range below 1 TeV [44]. The simplest SUSY model is the Minimal Supersymmetric Standard Model (MSSM), in which every SM particle has a superpartner, but the experiments confirm that these must have a larger mass than their SM counterpartners, and so, SUSY is also a broken symmetry. It resolves the Hierarchy problem, also provides dark matter candidates, and allows unification of all the fundamental forces at high energies [45].

- **Other New Physics Models.** Some other popular models include the existence of extradimensions [46, 47], or models in which the Higgs fields is not the source of electroweak symmetry breaking, such as Technicolor [48], where a new QCD-like interaction is introduced to break the electroweak symmetry by forming a condensate.

If the constrains of the SM are lifted, K, B and D decays and mixing are described by many more parameters than just the four CKM parameters and the W, Z, and quark masses. The most general effective lagrangian at lowest order contains around a hundred flavour changing operators, and the observables effects of interactions at the weak scale or above are encoded in their coefficients. The measurement of the magnitudes and phases of the CKM elements provide excellent sensitivity to new physics, so they should be measured as precisely as possible. The overconstraining measurements of CP asymmetries, mixing, semileptonic, and rare decays have started to severely constrain the magnitudes and phases of possible new physics contributions.

The LHC will be able to reduce the phase-space of some models, and hopefully, to see an evidence of the favoured New Physics model. When new particles are observed in the LHC, it will be important to know the flavour parameters as precisely as possible to understand the underlying physics.

Chapter 2

The LHCb experiment

The Large Hadron Collider (LHC) has been constructed at CERN and it is the particle accelerator with the highest energy in the center of mass (up to $\sqrt{s} = 14$ TeV). This is needed to study Physics beyond the Standard Model. Moreover, this accelerator has one of the highest luminosities and a production cross section of 500 μb (at $\sqrt{s} = 14$ TeV) [50]. One of the four LHC experiments, the LHCb, will analyse the $b\bar{b}$ quarks at this energy and will study the CP violation in the b system as well as rare B decays. LHC and LHCb will be described in the following section, giving a particular focus in the LHCb detector.

2.1 The Large Hadron Colider

The European Organisation for Nuclear Research known as CERN (Conseil Européenne pour la Recherche Nucléaire) is the worlds largest particle physics laboratory. It is located in Switzerland and France near the city of Genève. Together with laboratories like SLAC[1] or Fermilab[2] in the USA, DESY[3] in Germany and KEK[4] in Japan is one of the most important laboratories for particle physics. CERN operates with a yearly budget of 1.000 Millions swiss francs which is raised by its 20 european member states. Furthermore CERN employs $\sim$ 2500 people (out of which around 1000 are scientific personnel) and collaborates with $\sim$ 600 universities and institutes and nearly 8000 people world wide.

The Large Hadron Collider (LHC) is a circular high-energy proton-proton collider located at CERN, designed to operate at a center of mass energy of $\sqrt{s} = 14$ TeV and a luminosity of 10^{34} cm^{-2}s^{-1}. The main accelerator is located in a 100 m underground tunnel with a

[1] http://www.slac.stanford.edu
[2] http://www.fnal.gov
[3] http://www.desy.de
[4] http://www.kek.jp

circumference of 27 km, which was previously used to house the Large Electron Positron (LEP) accelerator.

At the LHC there are four main experiments, which are situated along the ring (see Figure 2.1). ATLAS (A Toroidal LHC ApparatuS) and CMS (Compact Muon Solenoid) are general purposes detectors, for the study of proton-proton collisions. They are dedicated to study high energy processes allowing the discovery of new particles, like the Higgs boson (in case it would exist according to the Standard Model) or new physics processes. ALICE (A Large Ion Collider Experiment), is dedicated to heavy ion physics, its primary aim is to study lead-lead collisions. ALICE will look for evidences of QCD bulk matter and quark-gluon plasma. Finally, the LHCb detector, is aimed to the study of CP violation and to the discovery of new physics via B meson decays.

FIGURE 2.1: Overview of the Large Hadron Collider, showing the positions of ATLAS, CMS, LHCb and ALICE.

Other smaller experiments are located in the LHC as well, like LHCf (LHC forward) and TOTEM (TOTal Elastic and diffractive cross section Measurement). LHCf studies energy distributions of particles in the very forward region. The LHCf detectors are placed at roughly ± 140 m from the ATLAS interaction point. TOTEM is designed to measure the total elastic and diffractive cross section as its name indicate, and shares point 5 with CMS.

The two 7 TeV counter-circulating proton beams of the LHC will contain 2808 bunches each, with a bunch-crossing frequency of 40 MHz. Each bunch contains approximately 10^{11} protons, resulting in an average of 600 million collisions per second[5].

To reach this energy, prior to the LHC injection, the protons go through several pre-accelerators. The CERN accelerator complex is a succession of particle accelerators that can reach increasingly higher energies. Each accelerator boosts the speed of a beam of particles, before injecting

[5]For high luminosity experiments (ATLAS and CMS)

it into the next one in the sequence, which functions as follows: protons are obtained by removing electrons from hydrogen atoms while hydrongen gas is bombarded with electrons from a hot cathode. The obtained protons are accelerated to 50 MeV by the CERN linear accelerator (Linac2) and injected through the Proton Synchrotron Booster (PSB), where they are accelerated up to 1.4 GeV. This protons are subsequently injected into the Proton Synchrotron (PS), where the initial energy is raised till 26 GeV, and into the Super Proton Synchrotron (SPS), from which they are transfered to the LHC at an energy of 450 GeV. In the LHC, the protons are accelerated during 20 minutes before reaching the maximum speed and energy (up to 7 TeV). Beams circulate for many hours inside the LHC beam pipes under normal operating conditions. The whole acceleration chain can be seen in Figure 2.2.

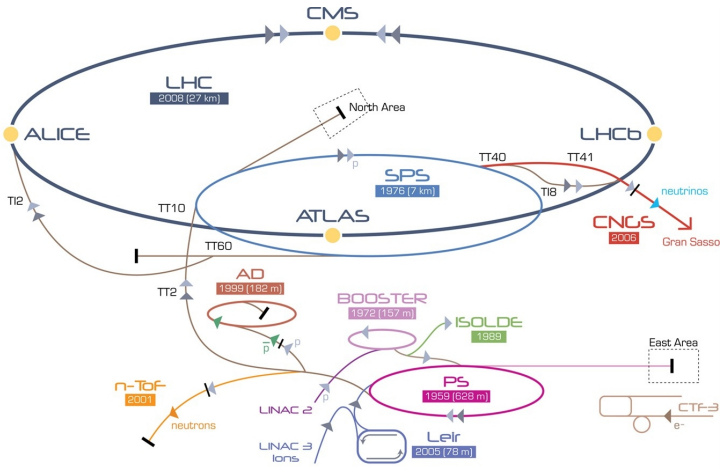


FIGURE 2.2: CERN acceleration complex. The protons are accelerated by the Linac2 and injected subsequently in the Booster, the Proton Synchrotron (PS), the Super Proton Synchrotron (SPS) and finally to the LHC where they can be accelerated till 7 TeV.

In the accelerator, particles circulate in vacuum tubes and are manipulated using electromagnetic devices: dipole magnets keep the particles in their nearly circular orbits, quadrupole magnets focus the beam, the accelerating cavities are electromagnetic resonators that accelerates particles and keep them at a constant energy thus compensating their energy loss. To avoid collisions with gas molecules, the internal pressure of the LHC is 10^{-13} atm, a regime known as ultrahigh vacuum.

There is a large variety of magnets in the LHC (around 9300), including dipoles, quadrupoles, sextupoles, octupoles, decapoles, etc. Each type of magnet contributes to optimize the particles trajectory. The dipoles of the LHC represented the most important technological challenge for the LHC design. The maximum energy that can be achieved is proportional to the strenght of the dipole field, given a specific accelerator circumference. The super-conducting dipole magnets of the LHC are able to provide a very high magnetic field of 8.3 T over their length. They use niobium-titanium (NbTi) cables, which becomes superconducting at 10 K (-263.2 °C).

The main role of the LHC cavities is to keep the 2808 proton bunches tightly bunched to ensure high luminosity at the collision points and hence, maximize the number of collisions. They also deliver radiofrequency (RF) power to the beam during the acceleration to the top energy. The LHC use eight cavities per beam, each one with an accelerating field of 5 MeV/m and they operate at 4.5 K (-268.7 °C). The LHC operates at the lower temperature of 1.9 K (-271.3°C), the temperature is reached by pumping superfluid helium into the magnet system.

In the LHC, as a result of the injection routine only 2808 bunch slots (of the 3564 available) can be filled. The LHCb experiment is 1/8 cycle out of phase plus 3 additional bunch periods due to the position of the interaction point relative to ATLAS and CMS. As a consequence only 2649 bunch crossings with both bunches can be filled.

2.2 The LHCb detector

The LHCb (Large Hadron Collider Beauty) experiment [49] is designed to study CP violation and rare decays in the B meson system in pp-collisions at the LHC machine with a center-of-mass energy up to $\sqrt{s} = 14$ TeV and a bunch-crossing frequency of 40 MHz. With the 2011 target luminosity of $\mathcal{L} = 3.5 \times 10^{32}$ cm^{-2}s^{-1}, and a measured $b\bar{b}$ cross section at $\sqrt{s} = 7$ TeV of $\sigma(pp \to b\bar{b}X) = (284 \pm 20 \pm 49)$ μb [51], a total amount of the order of 10^{12} $b\bar{b}$ pairs will be produced within the LHCb acceptance per year. It is intended to provide a better understanding of quark flavour physics and it will allow precise measurements of the different angles of the unitary triangles. The large production of B mesons will lead to the hope of revealing new physics beyond the Standard Model.

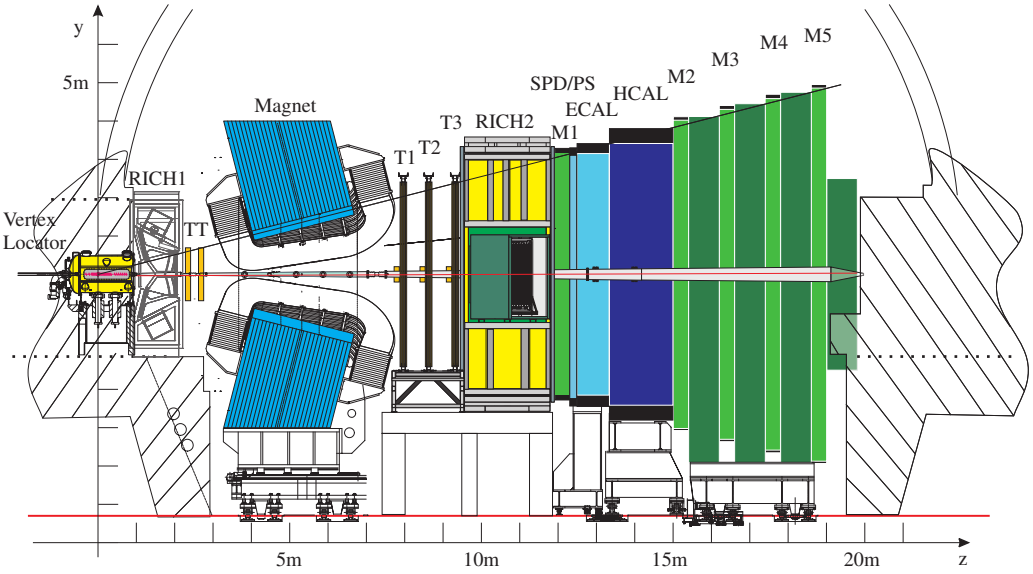


FIGURE 2.3: Schematic view of the LHCb detector with its subdetectors.

The LHCb does not cover 4π srad, it is a single arm spectrometer, in the forward direction, as shown in Figure 2.3. It focuses on the high rapidity region on one side of the interaction point, it covers the angles $10 < \theta < 300$ mrad ($1.9 < \eta < 4.9$) in the bending plane of the LHCb magnet. The design was chosen because the majority of $b\bar{b}$ pairs produced are expected to be in a narrow cone around the beam pipe, as shown in Figure 2.4. The production of $b\bar{b}$ quark pairs occur primarily through the interaction of gluons and sea quarks. About one third of the $b\bar{b}$ pairs are produced within the acceptance of LHCb.

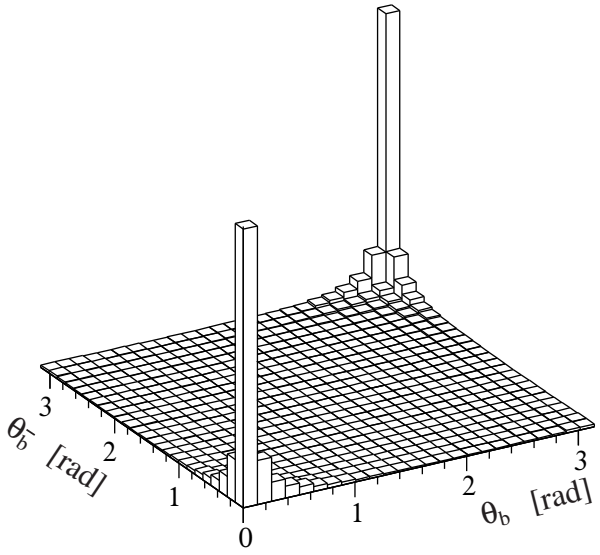


FIGURE 2.4: Azimuth angle distribution of the $b\bar{b}$ quark pair production at the LHC.

The LHCb operates at a lower luminosity than the LHC design luminosity. The LHC beams are defocused prior to collisions at Intersection Point 8, where the LHCb is located, to deliver a luminosity of 2×10^{32} cm^{-2}s^{-1}. This reduces the number of pp-interactions, operating in a regime where the probability of having one single interaction is maximized, making the identification of individual B mesons easier. Running at a lower luminosity simplifies the lifetime measurements and reduces the radiation damage to the detector. Figure 2.5 shows the probability for a number of interactions per bunch-crossings as a function of the luminosity.

The coordinate system of the experiment sets the z axis as the longitudinal to the beam (increasing z is called downstream, decreasing upstream), the y axis is the vertical one, and the x axis, the horizontal. The origin of the coordinate system is the interaction point, shifted by -11.25 m in the z direction from the center of the cavern, to make the best use of the space available.

Because of the high interaction rate, the low branching ratio of interesting B decay products and the low computer resources, an efficient rejection of data needs to take place. For this purpose a trigger system is built, consisting of a hardware based system (Level 0 - L0) and a pure software based trigger (High Level Trigger - HLT).

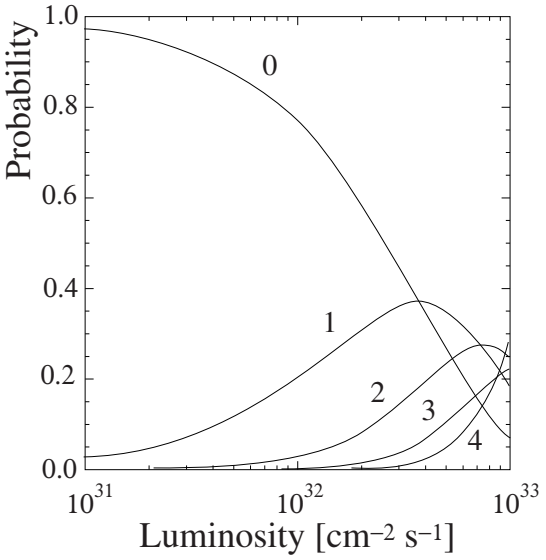


FIGURE 2.5: Probability of pp interaction per full bunch-crossing as a function of luminosity. Curves are shown for 0,1,2,3 and 4 pp-interactions.

The LHCb apparatus is divided into a number of subsystems: vertex locator, tracking and particle identification systems: Ring Imaging Cherenkov (RICH) detectors, calorimeters, and the muon system. Each of which is responsible for a particular part of the B meson measurement. The overall dimension of the LHCb detector is roughly 6 m $\times$ 5 m $\times$ 20 m. In the following sections the subdetectors, the trigger system, the software, the reconstruction and the data taking period are described.

2.3 The Beam Pipe

The beam pipe is needed to mantain the LHC vacuum. It is important that the particles created in the pp-collisions interact with it as little as possible. It is 19 m length and has four sections [52]. The first three sections, made of beryllium, account for 12 m, the fourth occupies a region where transparency is less critical and it is made of stainless stell. The beryllium is suitable because of its long radiation length, however it is a toxic, fragile and expensive material. The beam pipe joints the vertex locator (VELO) via a spherical metal shell made from an aluminium alloy. The joints between the four sections of the beam pipe are also made of aluminium.

2.4 The VELO

The VErtex LOcator, VELO [53, 54], is a fundamental component of the LHCb detector. It is responsible for the reconstruction of B mesons decay vertices, which are displaced secondary

vertices because of the relatively long lifetime of the b-hadrons. It provides a precise vertex location, required for the precise measurement of the lifetime ($\sim$ 40 fs) needed for many physics analysis. The VELO is also necessary for the measurement of the particle impact parameter and the reconstruction of the primary vertex. For $b\bar{b}$ events in which the b-hadron is produced in the LHC acceptance, the resolution of the primary vertex position depends on the number of tracks, but on average is 44 μm in the z-direction and 8 μm perpendicular to the beam [53]. Data from VELO is also used in the first and second level of the trigger. The VELO is designed to be retractable, moving the sensitive elements out of the path of the beam during the LHC injection, and repositioning them as close as possible during collisions. This is achieved by constructing the detector in two halves, which can be moved apart horizontally. The smallest possible distance of the detector from the beam axis is approximately 8 mm during physics runs [54].

A 1.8 m long section of the beam pipe around the interaction point has a large diameter to acommodate the vertex detector and its supporting mechanics. The VELO is made up of a series of silicon strip sensors, located within a secondary vacuum separated from the primary machine vacuum by a 100 μm aluminum foil. 21 sensor stations are placed transverse to the beam direction, as illustrated in Figure 2.6. The first 2 stations are used at the L0 trigger, as the Pile-Up sytem to count the number of interactions per event. The stations close to the interaction point enable to reconstruct tracks with angles up to 390 mrad. The downstream stations allows reconstruction of tracks down to 15 mrad.

FIGURE 2.6: The distribution of the VELO silicon sensor stations along the beam-pipe direction (left) and whole VELO detector (right).

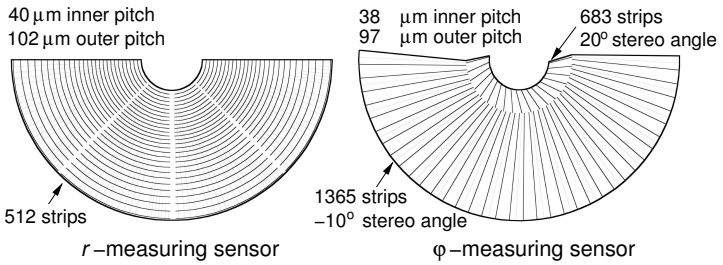


FIGURE 2.7: The strip layout for the R and Φ measuring detectors of the VELO detector.

The sensors are mounted in a total of 42 VELO modules, each module contains two detector disks to read orthogonal coordinates. One detector has circular strips measuring the R coordinate, while the other reads the Φ coordinate as shown in Figure 2.7. The R measuring sensors are segmented into 512 concentric semicircular strips, giving a strip pitch of 40 μm at the inner edge, increasing to 101.6 μm at the outer edge. The Φ measuring detector strips are divided into a inner region (683 strips) and an outer region (1365 strips), with a strip pitch in the inner and outer region of 35.5 μm and 95.6 μm respectively.

Using the z-position of the sensors along the lenght of the VELO, the hit information allows 3D tracks to be reconstructed. The sensors are designed to provide a spatial resolution, for 100 mrad tracks, of approximately 4 μm, in the region with the smallest strip pitch. The resolution in measuring track impact parameters is $\sim$ 15 μm at high transverse momenta ($\sim$ 10 GeV) and $\sim$ 300 μm at lower transverse momenta ($\sim$ 0.3 GeV).

2.5 The Magnet

When charged particles pass through a magnetic field, their trajectories are bent. The curvature of their trajectory allows us to determine the particle momentum and charge. The integrated magnetic field required to provide sufficient momentum resolution is around 4 Tm for tracks originating near the interaction point [55].

The LHCb uses a warm dipole magnet with the main component in the y direction. The warm magnet was found to have some advantatges over superconducting technology like its price, faster construction, faster ramping up of fields and potential for field inversions. It consist of two identical coils above and below the beam pipe, reflecting the angular acceptance of the detector, with a weight of 54 tons. The coils are situated within an iron yoke with a total weight of 1500 tons.

The total integrated field is on average 4 Tm with a deviation of 5% depending on the path through the magnet. It can produce a maximum magnetic field of just over 1 T. The momentum resolution that can be achieved, assuming a spatial resolution of the tracker of 200 μm, is better

than 0.5% for momenta up to 150 GeV. To minimize systematic effects due to charge asymmetric effects within the detector, the polarity of the magnet is periodically reversed.

2.6 The Tracking system

The LHCb tracking system consists of the VELO (see Section 2.4) and four tracking stations; the Tracker Turicensis (TT), situated upstream of the magnet, and three stations (T1-T3) downstream of the magnet. The downstream tracking stations use a combination of silicon microstrips and straw-tubes in different regions, called the Inner and the Outer Tracker respectively. Whereas the TT uses exclusively silicon technology. The purpose of the tracking system is to find charged particle tracks and measure their momenta and trajectory. The reconstruction efficiency for tracks traversing all the tracking stations is 94% and the expected momentum resolution for tracks in all tracking stations is $dp/p < 0.4\%$ with worse resolution for higher momentum tracks [52]. A layout of the tracking system can be seen in Figure 2.8.

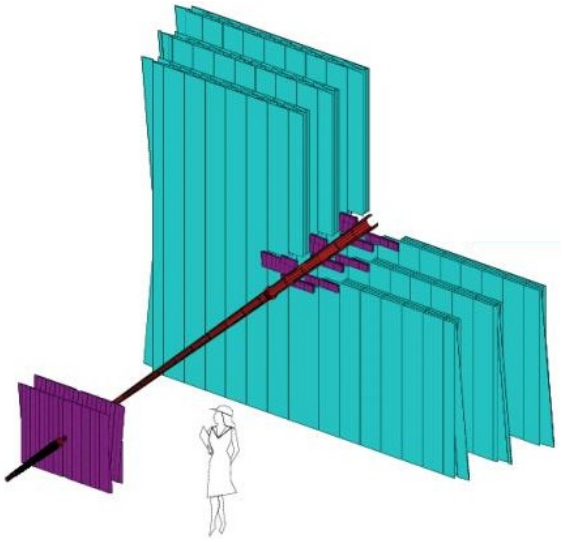


FIGURE 2.8: The LHCb tracking system and beam pipe. The Silicon Trackers (IT and TT) are depicted in purple and the Outer Tracker in blue.

2.6.1 The Tracker Turicensis

The TT station is responsible for providing tracking information on low-momentum particles that are swept out of the detector acceptance by the magnet. It also provides information for a simple momentum estimation which can be used at an early stage of a trigger algorithm. The TT station consist of 4 layers, arranged in two half sections separated by 30 cm along the beam axis enclosed in a single box providing electrical and thermal shielding. The layout is shown in Figure 2.9. The silicon sensors are 500 μm thick, 9.64 cm wide and 9.44 cm long. They carry

512 readout strips with a strip pitch of 183 μm. The sensors of the first and last layers (x layers) are disposed vertically while the ones of the second and third layers (u/v layers) are rotated a stereo angle of $\pm 5°$ respectively.

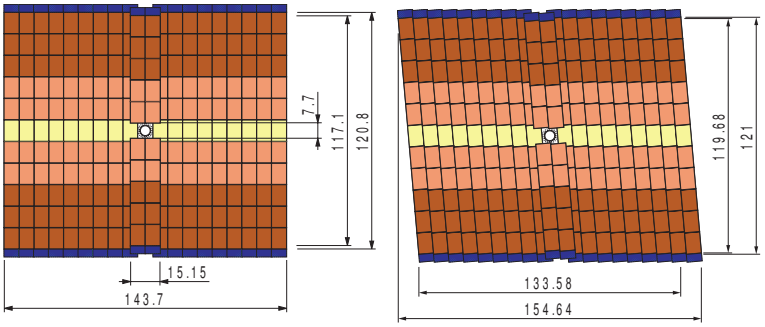


FIGURE 2.9: The layout in a TT-station. Lengths are in *cm*. The vertical (left) and rotated (right) readout strips, x and u planes are shown.

2.6.2 The Inner Tracker

The Inner Tracker (IT) [56] is located in the region closest to the beam pipe of the three downstream tracking stations (T1-T3). Each one is made up of four detector boxes as shown in Figure 2.10. Like the TT, each box is light-tight, electrically and thermally insulated and mantained at a temperature below 5°C. The side boxes consist of silicon-ladders like the TT ones but 320 μm thick, arranged with the same ($x-u-v-x$) orientation as those in the TT. The detector boxes above and below the beam pipe employ shorter ladders of one silicon sensor only. With this geometry, the detector design is adapted to the distribution of particle densities in the experiment.

2.6.3 The Outer Tracker

The Outer Tracker (OT) [57] is designed to track charged particles over a large acceptance range. It surronds the IT at each of the downstream stations in the region where particle fluxes and hence detector occupancy are much lower. It is assembled with an array of gas-tight straw tubes modules. Each module contains two staggered layers of straw tubes drift chambers, which makes a total of eight detection layers per station. The two double layers in the middle are placed also with the same stereo angle as the inner tracker, as it can be seen in Figure 2.8. The total area of the OT is 80.6 m^2 approximately and the total material in it corresponds to 9.6% of a radiation length.

A drift tube of the OT is assembled with a high carbon loaded polyimide film (Kapton© XC) wrapped with aluminium as the shell of the straw tube. The aluminium acts as shielding between

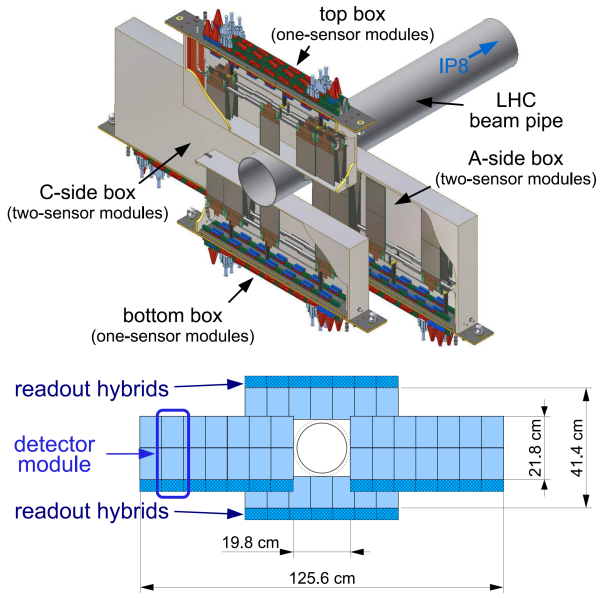


FIGURE 2.10: The layout of boxes in the IT, and of an x detector layer in the IT.

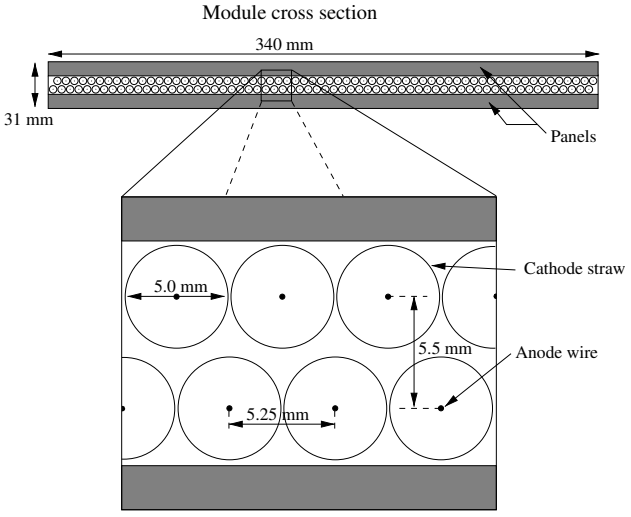


FIGURE 2.11: Staggering of straws within one layer of the Outer Tracker.

neighbouring channels. The central wire is a gold plated tungsten anode of 25 μm diameter. The tube itself has an inner diameter of 5.00 mm. The tubes are placed with a pitch of 5.25 mm in the modules as depicated in Figure 2.11. The drift gas used is a mixture of 70% Argon and 30% CO_2. A correction for the spillover has to be applied as the drift-time spectrum in the gas mixture has a width of 42 ns. The time-to-digital converter (TDC) readout provides drift-time with a precision corresponding to a spatial cell resolution around 200 μm and leads to a momentum resolution about 0.4% [58].

2.7 The RICH detectors

Particle Identification (PID) is crucial for an accurate reconstruction of B meson decays at LHCb, it is very important to distinguish between protons, pions and kaons, which are common final states in many B meson decays. In addition, identifying kaons is essential for the LHCb Flavour Tagging system, which determines the intial flavour of the B meson. Two Ring Imaging Cherenkov (RICH) detectors [59] are part of the LHCb PID system.

These two detectors, RICH1 and RICH2 (Figure 2.13), use Cherenkov radiation to obtain information on a particle's velocity. A charged particle radiates photons when traversing a dielectric medium faster than the light does. The charged particle emits radiation in a cone at a specific polar angle, the Cherenkov angle: $\theta_c = 1/n\beta$, where n is refractive index of the medium, and β is the ratio of the velocity of the particle to the speed of light. As only particles with velocities above a certain threshold $\beta_t = 1/n$ will radiate, multiple radiator materials with different index of refraction are needed to cover a wide velocity (momentum) range. LHCb uses two RICH detectors with different radiators to obtain sensitivity in a large momentum range. The RICH1, located before the magnet, uses silica aerogel and C_4F_{10}. The RICH2, situated behind the magnet uses CF_4. Figure 2.12 shows the Cherenkov angle for different types of particles, in different radiators, as a function of the momentum.

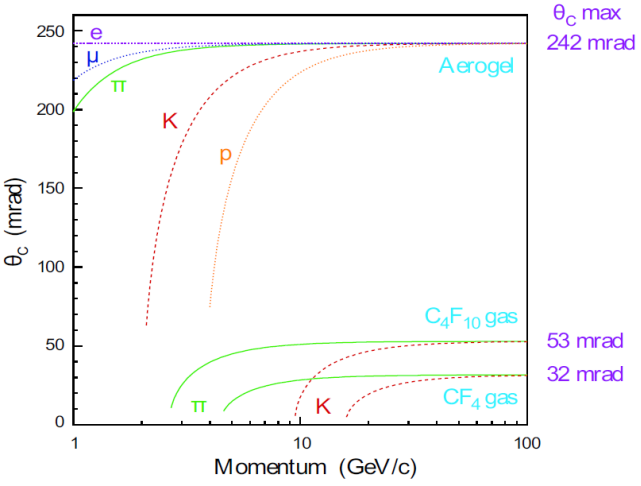
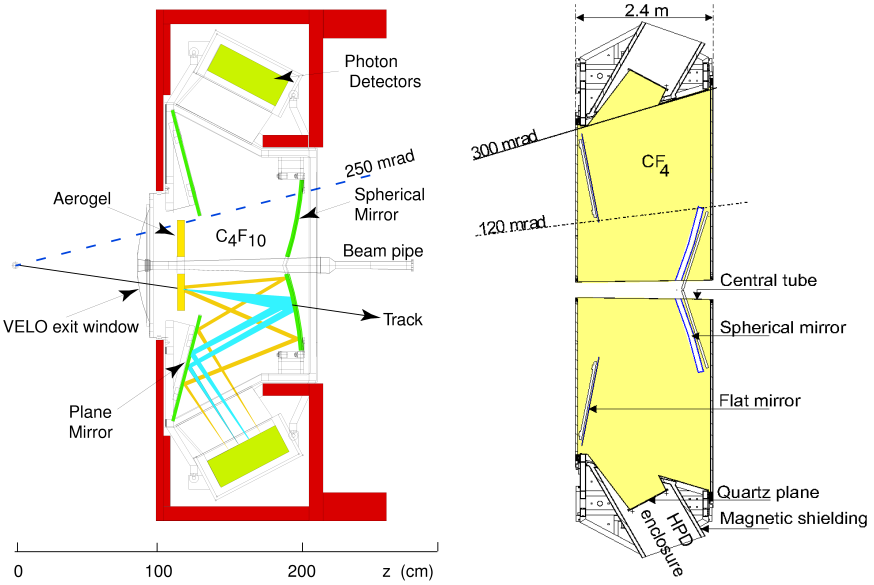


FIGURE 2.12: Cherenkov angle θ_C as a function of momentum for the different radiators used in LHCb and the different particles of interest. Aerogel and C_4F_{10} are used in RICH1, CF_4 in RICH2.

The RICH1 (see Figure 2.13(a)) covers the lower momentum region within the full LHCb, from 1 GeV/c to 60 GeV/c. The silica-based aerogel has a refractive index $n = 1.03$, and provides $\pi - K$ separation up to 10 GeV/c. The maximum θ_C for aerogel is 242 mrad, while it is only 53 mrad for the C_4F_{10}, that with a refractive index $n = 1.0014$ provides $\pi - K$ separation up to 60 GeV/c. To guide the light to the photon detectors, two set of mirrors are used: Spherical

mirrors from carbon fibre coated with aluminium and MgF_2 for focusing, and plane mirrors composed of a glass substrate with a coating of Al+SiO_2+HfO_2.

The RICH2 (see Figure 2.13(b)) situated downstream of the LHCb magnet covers an acceptance out of 120 mrad. It is designed to separate charged particles with a momentum between $\sim$ 15 GeV/c and 100 GeV/c. A single radiator medium, CF_4 gas is used, with a refractive index $n = 1.0005$ with a maximum $\theta_C = 32$ mrad. The layout of the RICH2 is similar of that of RICH1, rotated by 90 around th z-axis, so that the photodetectors planes are located at each side of the beam pipe, rather than above or below as in RICH1.



(a) RICH1 schematics (b) RICH2 schematics

FIGURE 2.13: The LHCb RICH detectors

In order to assign a PID hypotesis to a particle, a Cherenkov ring must be matched to a particle track, and for each mass hypotesis a likelihood is calculated. The required log likelihood difference (ΔLL) between the various PID hypotesis can be varied in order to tune the balance between efficiency and purity as required by the analysis.

2.8 The Calorimeters

The calorimetry system [60] of the LHCb is used for measuring the energy, position and shower shapes of electrons, photons and hadrons, and it uses this information to identify them. A calorimeter detects an incident particle by absorbing all of its energy. Photons and electrons interact via electromagnetic processes while hadrons interact via nuclear interaction processes. These measurements are also required for various trigger algorithms (for which the transverse

energy is measured) and for the offline event analysis. An efficient and fast response is needed for the inclusion in the L0 trigger.

The calorimetry is located downstream the RICH2, and, in downstream direction. Its polar acceptance goes from 30 mrad to 300 mrad horizontally and 250 mrad vertically. The calorimeter consist of: a Scintillator Pad Detector (SPD), to differentiate between charged and neutral particles; the Pre-shower (PS), which provides additional information on the shower shape and is used to distinguish between charged pions and electrons; an electromagnetic calorimeter (ECAL), which enable to reconstruction of photons, electrons and neutral pions; and the hadronic calorimeter (HCAL).

The four elements of the calorimetry system employ similar technologies, i.e. scintillators coupled to wavelength-shifting fibres (WLS) read out by fast photodetectors, with the advantage that the front-end electronics is largely unified for the whole calorimeter. Figure 2.14 shows the performance of the LHCb calorimeter in front of different kind of particles.

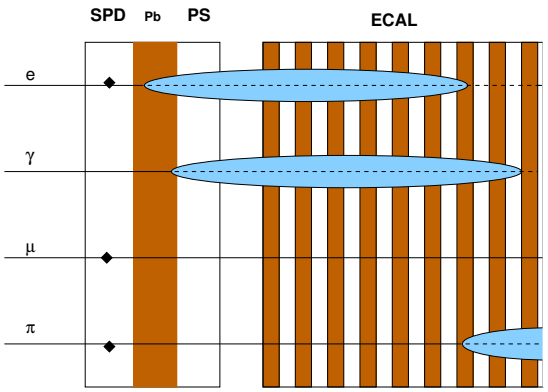


FIGURE 2.14: Schematic view of electromagnetic showers generated by electrons and photons at PS and ECAL systems. Also muon and pion behaviour is shown. Charged particles deposit energy in the SPD by ionization as represented with the sign ♦.

2.8.1 Scintillator Pad Detector

The Scintillator Pad Detector (SPD) is meant to complete the calorimeter system performing the electron and photon separation at the L0 trigger [61]. It also aids π^0 identification, which decay to a pair of photons. The SPD design consists in a matrix of plastic scintillator pads with an optical fibre coiled inside, which is used to collect and guide the scintillator light to a photomultiplier. The scintillators have an area of 7.6×6.2 m^2, and are highly segmented. Due to the large variation in hit densities over the surface, it is divided into three regions with different granularities. The inner section, closest to the beam pipe, is comprised of 1472 cells of size approximately 4×4 cm^2, the middle section has 1792 cells 6×6 cm^2, and the outer section is the largest one and contains 2688 cells 12×12 cm^2[60].

The scintillator light is transported via wavelength-shifting (WLS) fibers, and measured by 64 channel Multi-Anode Photo Multiplier Tubes (MAPMT) placed at the top and the bottom of the detector. While the size of the signal in each cell of the PS must be recorded, the SPD output is a single bit output telling whether a cell has been hit or not.

Charged particles crossing scintillators produce ionization, while neutrals contribute little. When a charged particle crosses the scintillator tile it leaves an average of 2.85 MeV in the scintillator, and corresponds to an average of $\sim$ 25 photoelectrons in the PMT. To discriminate between neutral and charged particles a threshold can be applied for the energy deposited in a SPD cell. A cut on 0.7×2.85 MeV gives a 100% efficiency for electrons and reject most of the photons [62].

The proper performance of the SPD, and thus of the L0 trigger, depends on a good calibration of the SPD. Different strategies had been developed for this purpose. My personal contribution was to develop a calibration strategy. This strategy consist on thresholds scans with an analytical fit to the data, collected with a special L0 trigger configuration based on a single cut on the calorimeter energy. Assuming standard luminosity, scanning the region from 0.5 to 4.5 MeV, using 20 steps of 0.2 MeV, and spending 30 seconds taking data in each step, one should be able to extract stable and unbiased measurements of the number of photoelectrons and deposited energy peak position at a 2% error level. The total amount of time needed to calibrate the SPD should not be longer than 20 minutes [63]. The low luminosity during the first days of the LHCb did not allow the SPD calibration with the mentioned method. Instead, it was applied a calibration method based on track extrapolation, which is detailed in [64].

2.8.2 Pre-Shower

The main purposes of the preshower (PS) detector [61] are the electron/hadron (pions particularity) separation by providing additional information on the shower shape, as well as the correction for the ECAL energy of electromagnetic showers that begin before the calorimeter. To achieve the first purpose a 12 mm lead converter plane after the SPD is used. A scintillator plane, similar to SPD, is used to measure the beginning of the shower. The SPD/PS system is build achieving a one-to-one correspondence between the sections of the SPD and the PS. The electron/hadron separation is based in the fact that the electrons in contact with the lead converter produce a shower of secondary particles reaching the scintillator and inducing a signal much larger than a typical signal from a charged pion. For a threshold of four Minimum Ionising Particles (MIP's), pion rejection factors of 99.6%, 99.6% and 99.7% with electron retentions of 91%, 92% and 97% are achieved for 10, 20 and 50 GeV/c particle momentum.

2.8.3 Electromagnetic CALorimeter

The main aims of the electromagnetic calorimeter (ECAL) are: a) to provide electron identification at different trigger levels and at the offline reconstruction stage; b) to measure the energy of electromagnetic particles (i.e. electrons and photons); and c) to provide the π^0 reconstruction. The ECAL uses shashlik technology, that is a sampling of alternate layers of lead and scintillator structure with plastic WLS readout fibres. It has the same segmentation in the $x-y$ plane as the SPD/PS, although each cell is approximately 1.5% bigger. The ECAL begins 12.49 m downstream of the interaction point, and extendes for 835 mm. The detector is built as two separate halves from modules that are fixed to a surrounding frame. Modules in different sections (inner, middle, outer) contain a different number of readout cells.

The ECAL energy resolution is a function of the particle energy, and it is designed to have a $\sigma(E)/E = 10\%/\sqrt{E} \oplus 1\%$ where the energy is measured in GeV [60]. This, together with the preshower and tracking information, provides sufficient electron/hadron separation at the trigger level as well as at the reconstruction stage.

2.8.4 The Hadronic CALorimeter

The HCAL is responsible for measuring the energy of the hadrons, the main purpose is to provide data for the L0 hadron trigger. The HCAL begins at a distance of 13.33 m from the interaction point, and is 1.65 m deep. The required energy resolution is moderate: $\sigma(E)/E = 80\%/\sqrt{E} \oplus 10\%$. However the detector has to be fast in order to measure the particles transverse energy at 40 MHz.

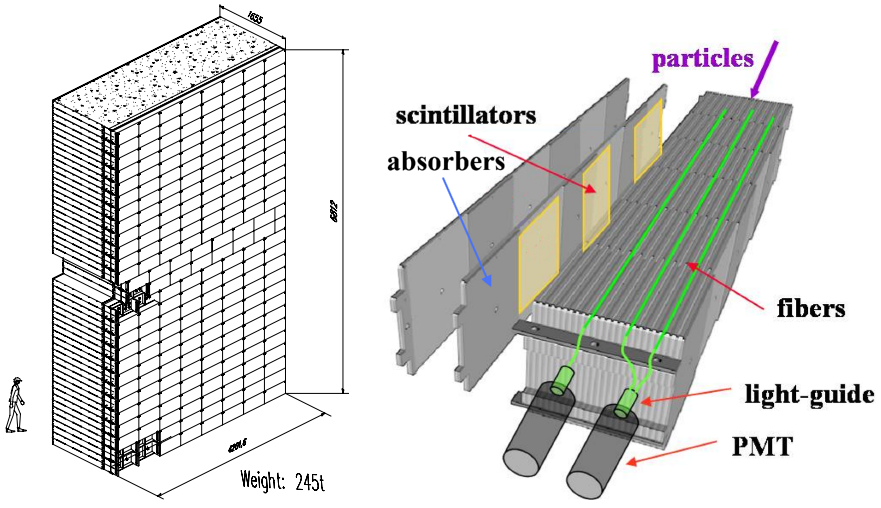


FIGURE 2.15: View of the LHCb Hadronic calorimeter (left) and a HCAL submodule structure (right).

The HCAL is made of steel as absorber and scintillator tiles as active material. The tiles are oriented so that they are parallel to the beam axis. In the lateral direction the tiles are interspersed with 1 cm of iron. The steel plates have a thickness of 16 mm with an average of 4 mm of scintillator in between. The HCAL is also made of 52 modules, divided into eight supermodules, and is segmented in two different sizes, the inner section with a width of 131.3 mm and the outer section, with a width of 262.6 mm [60]. Figure 2.15 shows its structure. Like the ECAL, the scintillator light is extracted via WLS fibres, and measured by PMTs.

2.9 The Muon system

The final LHCb subdetector is the Muon system [65], which is responsible for the identification, p_T measurement and triggering of the muons. It consists of five stations, M1-M5, the first one upstream the calorimeter, and the rest downstream, as shown in Figure 2.16. The muon identification is of fundamental importance to the LHCb experiment, as muons are present in the final states of several possible CP violation decays as well as rare decays. In addition, muons of semileptonic B decays can be used as a flavour tag, providing the initial flavour of the neutral B meson. As the muon information is used in the L0 trigger, a fast response is mandatory.

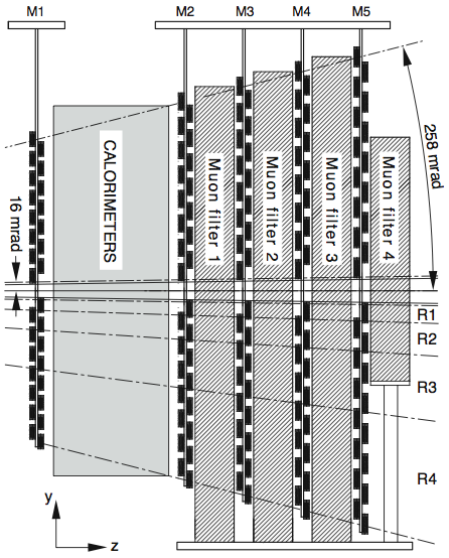


FIGURE 2.16: The LHCb muon system.

The first station provides an increased precision in the p_T measurement at trigger level. The remaining stations are separated by muons filter, which are 80 cm thick iron absorbers. M1 - M3 measure the muon p_T with a resolution of 20%. The M4 - M5, with a lower spatial resolution, are used to select highly penetrating muons. The muon system comprises 20 radiation lenghts in total, with a minimum momentum of 6 GeV required for a muon to traverse all five stations.

A combination of Multi Wire Proportional Chambers (MWPC), and Gas Electron Multiplier (GEM) detectors are used.

The muon stations are divided into four regions, R1 to R4, with a linear dimension ratio of 1:2:4:8 in the x-axis. Such ratio gives regions with approximately equal particle flux and occupancy. The MWPC are used in all regions except in the R1 region of the M1, where the MWPC would age to quickly because of the high particle flux, so GEM detectors are used. The size of a muon station ranges from 7.7×6.4 m^2 to 11.9×9.9 m^2 (M1 to M5).

2.10 The Trigger system

At a luminosity of 3.5×10^{32} cm^{-2}s^{-1} at a bunch crossing frequency of 40 MHz only 12 - 15 MHz of crossings have visible interactions. As it is not possible to write the event data from all collision to tape, a trigger system is used [66], selecting events that are stored for offline analysis, and so it must select pp interactions containing B and C mesons within the LHCb acceptance, while keeping the total rate to be written out down to a manageable rate, rejecting as many uninteresting background events as possible. The running conditions in 2010 and 2011 have changed with respect to those expected from the trigger design, thus, the trigger and store working conditions changed too. The output of the trigger in 2011 has been 3 kHz of clean samples of b and c decays.

There are two levels of trigger: the Level-0 (L0), consisting of electronics operating synchronously with the LHC machine, reducing the rate to 1 MHz; and the High Level Trigger (HLT), which reduces the rate down to 3 kHz asynchronously, using a processor farm.

To collect data from the front-end (FE) electronics of each subdetector as well as to control and monitor the experiment, an online data acquisition system (DAQ) and an experiment control system (ECS) had been developed [67]. After the L0 the DAQ provides the data from the FE to the HLT, running on the processor farm.

2.10.1 Level-0 Trigger

The L0 trigger comprises three components: the VELO Pile-Up system, the calorimeters, and the muon system. The Pile-Up veto attempts to find the number of primary pp interactions vetoing the events with a high number of pp interactions where the tendency would be to trigger in combinatorics rather than in genuine B decays. The L0 also attempts to reconstruct the highest E_T hadron, electron and photon clusters in the Calorimeters, and the two highest p_T muons in each quadrant of the Muon system. The information of each of these subdetectors is sent to the Level-0 Decision Unit (L0DU) in order to be combined to produce the overall

trigger decision. The fully synchronous L0-trigger, with a fixed latency of 4 μs, is implemented in custom boards and reduces the rate from 40 MHz to 1 MHz with efficiencies of approximately 90%, 70% and 50%, which are achieved for events with muons, hadrons and photons respectively.

2.10.2 The High Level Trigger

The high level trigger *HLT* is implemented in software that runs on an Event-Filter Farm (EFF) of about 2000 nodes. The HLT has access to all the data from each event, but given the 1 MHz data rate and the limited computing power it aims to use only part of the information available to make the decision of considering an event or not. The HLT is implemented in two stages, *HLT1* and *HLT2*. The final output of the HLT is sent at an event rate of 3 kHz to mass storage for offline analysis.

The HLT1 is designed to minimize the impact of varying running conditions on its performances. It is based around a single track trigger, which searches for a single track with a high momentum, a large impact parameter to all primary vertices in the event, and a good track quality. In addition, lifetime unbiased muon and electron triggers are used to confirm the L0 trigger decision. The $\sim$ 50 kHz selected by the HLT1 are then processed by the HLT2, which runs a full reconstruction, similar to the offline analysis. The HLT2 consist of inclusive and exclusive selections that cover the various decay channels required by the LHCb physics goals. Since the HLT is implemented in software, it is quite flexible, allowing the triggering strategy to be changed if necessary during the running of the experiment.

2.11 The LHCb software

All the information available from the detector is processed by dedicated software to be used for physics analysis. The LHCb software packages are based on an architecture known as Gaudi [68]. It supports both online and offline software. The framework is designed to accomodate changes in the requirements and detector technologies. The LHCb software was developed in C++ and its main applications are the following:

- **Gauss**[69] for event generation and detector simulation, which can be interfaced with a number of specialist software packages, such as Pythia, Herwig and EvtGen for event generation, or Geant4 for detector simulation.
- **Boole** [70], used to digitalize the data from Gauss to model the detector response.
- **Moore** [71], is used to run the HLT in the Online System, processing real data from the LHCb DAQ system, or offline starting from real data or from the output of Boole.

- **Brunel** [72] takes care of the reconstruction from both experimental data from the LHCb DAQ and from the output of Boole.
- The physics analysis is carried out by **DaVinci** [73] application.
- Event and detector visualisation is performed with **Panoramix** [74].

2.11.1 Reconstruction and analysis

The data recorded by the LHCb and the simulated events contain the same information. The processing of data starts with the RAW files obtained from data or Monte Carlo simulation (RAWmc). The Brunel package performs track finding and fitting, creates clusters from the calorimeter information and runs the particle identification algorithms to reconstruct the so called "proto-particles".

The particle trajectory is estimated by the track reconstruction program, which combines the hits in the VELO and in the tracking stations, taking into account the deviation due to magnetic field. Depending on their trajectory through the tracking system tracks are defined as:

Long tracks cross the full tracking system from the VELO to the T stations. These tracks have the most precise momentum and therefore are the most commonly used for physics analyses. The reconstruction efficiency from a particle with a momentum larger than 10 GeV/c is on average $\sim 94\%$.

Upstream tracks transverse only the VELO and the TT stations. These are low momentum tracks that are bent out of the LHCb acceptance for the magnetic field.

Downstream tracks only transverse TT and T stations. They are relevant in the reconstruction of long-lived particles which decays outside the VELO.

VELO tracks are measured only in the VELO and are useful for primary vertex reconstruction.

The particle identification system uses the RICH detectors, the calorimeter system and the muon detector to associate to each proto-particle a particle identification likelihood for each of the long-lived mass hypotheses (pion, kaon, muon and proton). For that, the different particle identification contributions are combined into a log-likelihood difference between a given PID hypothesis and the pion hypothesis. This creates a particle for each charged track or neutral cluster.

Brunel also identifies the primary vertices in an event. In case of simulated events, Brunel associates the reconstructed proto-particles with the generated (simulated) particle. The event reconstruction of the RAW files is stored in the Data Sumary Tapes (DST), which are the outputs of the stripping selections - which filters the events for physical analyses - and contains the desired information to perform physics analysis.

The physics analysis is performed by the DaVinci package. It is used to reconstruct particle decay chains (like $B_s^0 \to J/\psi\phi$) and to study their properties. These particles can then be used to create other particles, with associated properties that can be calculated from the kinematics. By requiring particle properties to lie in certain ranges, they can be filtered to select those most likely to match a decay chain.

A distributed computing system, called GRID [75], has been developed for the analysis of LHC data, and allows users to run analysis jobs remotely, wherever processing power is available. The job submission manager package was developed by the LHCb and ATLAS collaborations, and is called Ganga [76]. The computing model of the LHCb consist of a central site (CERN) responsible for distributing RAW data to the Tier-1 centers (CERN, CNAF, FZK, IN2P3, NIKHEF, PIC and RAL) where the DSTs will be produced and analysed. The Tier-2 centers are primarily Monte Carlo production sites.

2.12 Data taking period

The first LHC collisions took place on the 23rd November 2009 with an energy of 450 GeV per beam. But it was on 30th March 2010 when the first proton-proton collisions at 3.5 TeV per beam signaled the arrival of the largest hadron collider in the world, the LHC era had started. The 3.5 TeV collisions continued until the end of October 2010, when the first heavy ion collisions took place. By the end of 2010 the LHC had delivered 42.15 pb^{-1} of integrated luminosity of which LHCb recorded 37.66 pb^{-1} ($89\% efficiency$). The LHCb detector performed well during this first high energy run, the configuration of the LHC machine was set to maximize the luminosity, the number of bunches increased to 344 with 1.1×10^{11} protons per bunch resulting in an instantaneous luminosity of 1.6×10^{32} $\text{cm}^{-2}\text{s}^{-1}$. On the other hand, the number of visible collisions per bunch crossing ($\mu \sim 2.5$) increased above the design value ($\mu = 0.74$). The majority of the luminosity was recorded with a peak μ greater than 1.5.

A second period of 3.5 TeV proton-proton collisions started on 12th April 2011. During the 2011 LHCb physics run the μ-value was still higher than the design value ($\mu = 1.5$). As a consequence the number of tracks and vertices per event increased, creating a challenging environment for the trigger and reconstruction systems. At the end of 2011 the LHC delivered ~ 1.2 fb^{-1} to LHCb, of which ~ 1.1 fb^{-1} have been recorded. This corresponds to an efficiency of 91 %. Data has been collected with up to 1300 colliding bunches per beam, and a target luminosity of 3.5×10^{32} $\text{cm}^{-2}\text{s}^{-1}$. For 2012 a further run is scheduled at 4 TeV per beam before a long shutdown to commission the LHC to its design energy of 7 TeV per beam.

Possible New Physics is expected in a wide range of CP measurements and rare decays where LHCb will be able to perform precise measurements. There are some key measurements for the

LHCb detector: the measurement of the mixing-induced CP violation in $B_s^0 \to J/\psi\phi$, oscillation measurements, analysis of radiative decays and rare decays as $B_s \to \mu^+\mu^-$, etc. To perform CP violation measurements and for the measurement of $B^0\bar{B}^0$ oscillation it is crucial to know the initial flavour of the reconstructed B^0 (or $\bar{B}^0$) and B_s^0 (or $\bar{B}_s^0$) mesons, to establish whether the meson contained a b or a $\bar{b}$ quark.

Chapter 3

Flavour Tagging

To observe B meson oscillations and to make precise CP violation measurements, it is crucial to have a good performance in the determination of the flavour of the reconstructed B meson. This procedure is known as *Flavour Tagging* and is performed at LHCb by means of several algorithms. In this Chapter it is explained in detail the opposite side and same side flavour tagging algorithms. We also report on the process of obtaining the tagging decision and mistag probability.

3.1 Introduction

In LHCb, B mesons are mainly produced as $b\bar{b}$ pairs. One of the two mesons decays, called signal B (B_{sig}), is completely reconstructed, while the other, called tagging B (B_{tag}) or opposite B (B_{opp}) decays inclusively and is used to tag the initial flavour of the signal B. The identification of the initial flavour of reconstructed B^0 and B_s^0 mesons is necessary for most of the measurements of flavour or CP asymmetries to establish whether the meson contained a b or a $\bar{b}$ quark. This procedure is known as *flavour tagging* and it is performed at LHCb by means of different flavour tagging algorithms, known as *taggers*, which can exploit the "same side (SS)" and the "opposite side (OS)" with respect to the signal B as illustrated on Figure 3.1. To see flavour oscillations or CP violating asymmetries, apart from a sufficient large data sample and an adequate reconstruction precision, a good flavour tagging is mandatory.

The flavour tagging in LHCb has 4 opposite side taggers (muon, electron, kaon and vertex charge tagger) and 2 same side taggers (SS pion and SS kaon). A decision is made for each tagger establishing whether the signal meson contained a b or a $\bar{b}$ quark. The probability of the tag decision to be correct is also estimated by using several kinematic properties of the tagger and the event itself. When more than one tagger is available per event, these probabilities are

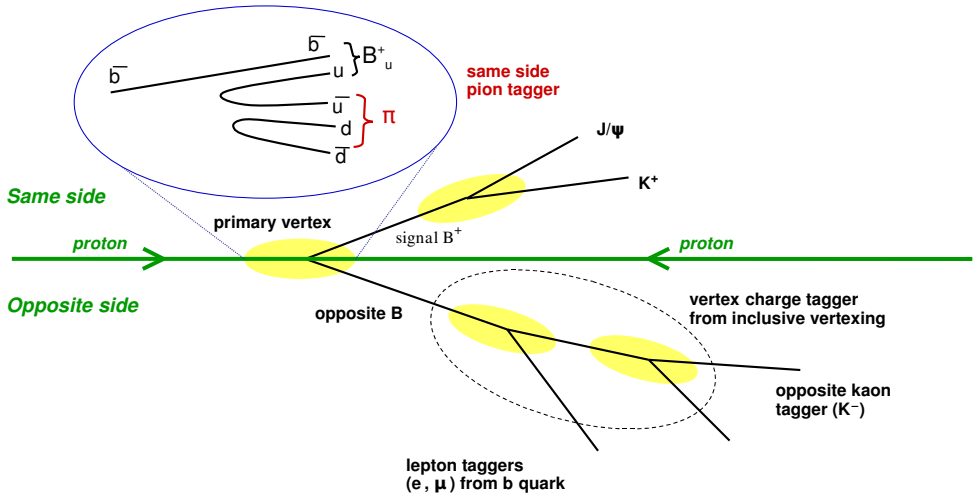


FIGURE 3.1: Schematic representation of the different sources of information available to tag the initial flavour of a signal B candidate, here $B^+ \to J/\psi K^+$. The SS and OS (which can be any b hadron) are shown.

combined into a single decision per event to provide an estimated mistag probability ω which can be used in the CP-fits to calculate CP asymmetries.

3.1.1 Definitions

The tagging procedure is never perfect and the observable asymmetry will always be smaller than the true asymmetry, introducing a dilution term as it can be seen in Figure 3.2.

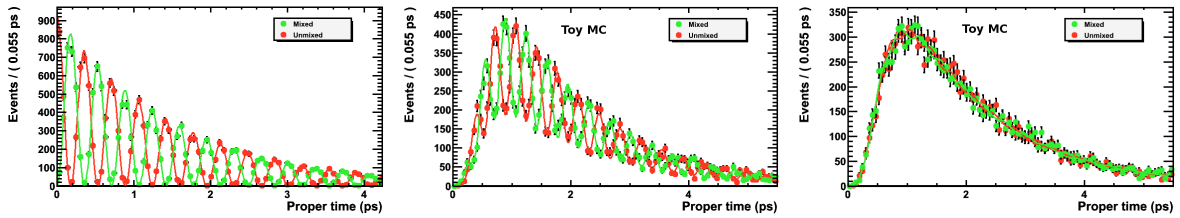


FIGURE 3.2: Simulation of B_s^0 mixing. Red line shows non-oscillated B mesons while green line shows the oscillated B mesons. Left: Ideal measurement, with no resolution or detector effects. Middle: Realistic acceptance function and realistic mistag. Right: Realistic time resolution added to the realistic acceptance and mistag.

The dilution can be due either to acceptance and time resolution effects, to the introduction of background in the selection, and to inefficiencies in the tagging algorithms. These last inefficiencies are mainly caused by:

- mixing of the B_{opp} (if B^0 or B_s^0),
- wrong sign lepton tags from semileptonic decays through a charm $b \to c \to l$,

- other particles from the underlying event used as tagging particles,
- wrong tags due to particle misidentification

All the reconstructed B meson decays can be divided into correctly tagged (R), uncorrectly tagged (W) or untagged events (U). The fraction of events with a tagging decision (d) available is called tagging efficiency ϵ_{tag} and is calculated as:

$$\epsilon_{tag} = \frac{R+W}{R+W+U} \tag{3.1}$$

The fraction of events of which a wrong decision was taken is called wrong tag fraction ω and is calculated as:

$$\omega = \frac{W}{R+W} \tag{3.2}$$

The measurement of the mistag fraction is possible on flavour-specific decay channels. For charged mesons this is obtained by directly comparing the flavour of the reconstructed meson with the tagging decision, while for neutral mesons this is done by fitting the B^0 flavour oscillation as a function of the proper time. For non flavour-specific channels you have to trust the tagging algorithms response to estimate an ω event-per-event. To obtain a reliable wrong tag fraction, the ω has to be correctly calibrated, as explained in Chapter 4.

The tagging efficiency and wrong tag fraction are directly related to the statistical error on the CP asymmetry. The measured rate of a B decay to a final state f ($B \to f$) and its CP conjugate can be writen as:

$$\begin{aligned} \Gamma^m(t) &= (1-\omega)\Gamma(t) + \omega\overline{\Gamma}(t) \\ \overline{\Gamma^m}(t) &= \omega\Gamma(t) + (1-\omega)\overline{\Gamma}(t) \end{aligned} \tag{3.3}$$

So the measured time dependent asymmetry ($\mathcal{A}_m$) can be computed as:

$$\mathcal{A}_m = \frac{\overline{\Gamma^m}(t) - \Gamma^m(t)}{\overline{\Gamma^m}(t) + \Gamma^m(t)} = (1-2\omega)\mathcal{A}(t) = D\mathcal{A}(t) \tag{3.4}$$

where $D = (1-2\omega)$ is the dilution term. A perfect tagging gives $D = 1$, while random tagging ($\omega = 50\%$) gives $D = 0$. A negative dilution means that the flavour tagging anti-tags the b. The dilution decreases the measured oscillation amplitudes. The real asymmetry $\mathcal{A}$ and the statistical error on the asymmetry is given by

$$\mathcal{A} = \frac{\mathcal{A}_m}{1-2\omega} \qquad \Rightarrow \qquad \sigma_{\mathcal{A}} = \frac{\sigma_{\mathcal{A}_m}}{1-2\omega} \tag{3.5}$$

So, as $1 - \mathcal{A}_m^2 = \frac{4\Gamma^m\overline{\Gamma^m}}{(\Gamma^m+\overline{\Gamma^m})^2}$, using quadratic propagation of the errors for Eq. 3.4, the error in the measured asymmetry is given by

$$\sigma^2_{\mathcal{A}_m} = \frac{4\Gamma^m\overline{\Gamma^m}}{(\Gamma^m+\overline{\Gamma^m})^3} = \frac{1-\mathcal{A}_m^2}{\overline{\Gamma^m}(t)+\Gamma^m(t)} = \frac{1-\mathcal{A}_m^2}{N^m} = \frac{1-\mathcal{A}_m^2}{\epsilon_{tag}N} \tag{3.6}$$

where N^m are the number of events where the initial flavour is known and N is the total number of events. Then the error on the true asymmetry is given by

$$\sigma_{\mathcal{A}} = \frac{\sqrt{1-\mathcal{A}_m^2}}{\sqrt{\epsilon_{tag}N}(1-2\omega)} \tag{3.7}$$

which shows that to minimize the statistical error we need to maximize the effective tagging efficiency ϵ_{eff}, also called tagging power, that is defined as:

$$\epsilon_{\mathit{eff}} = \epsilon_{tag}D^2 = \epsilon_{tag}(1-2\omega)^2 \tag{3.8}$$

As an example, in the case of $B^0 \to J/\psi K_S$, the CP measured asymmetry can be written as,

$$\mathcal{A}_m^{B^0\to J/\psi K_S} = (1-2\omega)\sin(2\beta)\sin(\Delta m_d t) \tag{3.9}$$

and the statistical error on the $\sin(2\beta)$ measurement has the form

$$\sigma_{\sin(2\beta)} \propto \frac{\sqrt{1-\mathcal{A}_m^2}}{\sqrt{\epsilon_{tag}N}(1-2\omega)} \propto \frac{1}{\sqrt{\epsilon_{\mathit{eff}}N}} \tag{3.10}$$

As mentioned before, the sensitivity of the measured asymmetry is directly related to the effective tagging efficiency ϵ_{eff}, or tagging power, which directly depends on the tagging efficiency (ϵ_{tag}) and on the wrong tag fraction (ω). These are the three important parameters in the tagging algorithms and are defined in Equations 3.8, 3.1 and 3.2 respectively. The different flavour tagging algorithms were designed with the aim of maximizing the ϵ_{eff} obtaining the best possible tagging performance.

3.2 Flavour Tagging Algorithms

The algorithms used to tag the initial flavour of the signal B are written within the LHCb analysis framework called DaVinci, and can be found inside the *Flavour Tagging* (FT) package. The tagging algorithms, also referred to as *taggers*, use the charge of the tagging particle, or

the vertex charge in case of the inclusive reconstruction of the B_{opp}, to give a tagging decision d. They also estimate the probability of the tagging decision to be correct $p = 1 - \omega$ and the mistag probability ω using several kinematic and geometrical properties of the tagger and of the event itself. To do that, the taggers use all the information available in the event in order to obtain an optimal statistical precision.

The tagging algorithms were initially designed on Monte Carlo (MC) simulations [77, 78]. Since then, the LHCb tagging algorithms have been adapted, redesigned and reoptimized using the new MC productions and data. In the following sections we will consider the last iteration of this process which was performed on MC10 data produced with DaVinci v28r4. The $B^+ \to J/\psi K^+$ channel was specifically used for that purpose although any other channel could have been used because the tuning of the algorithms is channel-independent. This statement is clearly understood for OS taggers as the B_{opp} is completly independent of the B_{sig}. For the SS kaon or SS pion taggers, the selection of the B_{sig} could introduce some bias in the spectrum of the SS tagging particle, although no large differences are expected.

In order to develop the tagging algorithms, a *Flavour Tagging Checker* (FTC) software package was created. The FTC package reproduces the output of the FT package but it also allows us to modify the tagging algorithms to study the performances. It is used to optimize and calibrate the tagging performance as well as to perform the training of the neural nets which provide the mistag estimation event-per-event. In particular, the FTC package is used to find the relevant variables which distinguish better between right and wrong tagged events as well as to set the cuts on the tagging candidate which maximize the tagging performance. It counts the right and wrong tagged events by comparing the tagging decision with the flavour of the reconstructed signal B meson, and compute the effective tagging efficiency, which is the value we want to maximize.

Although the internal structure of the taggers is based on the MC10 studies, the data collected by LHCb since 2010 allowed us to reoptimize and calibrate the taggers on data. These procedures are described in Chapter 4. The cuts inside the FT package can be easily changed via options and adapted to new future optimizations and calibrations.

3.2.1 Tagging candidates

Not all the particles in the event can be used for tagging purposes, only a subsample of them are considered as tagging candidates. To select a particle as a tagger, a preselection of tagging candidates is done. Only charged particles associated to "good quality" ($\chi^2/d.o.f < 5$) *long*[1]

[1]Require hits in the Velo and in T-stations.

and *upstream*[2] tracks are considered as tagger candidates. The particles are required to be outside of a cone of 5 mrad (labeled as $\Delta\phi$) formed by any daughter of the signal B. A maximum momentum (200 GeV/c) and transverse momentum (10 GeV/c) are requested as well as a minimum momentum of 2 GeV/c to exclude low momentum particles coming from the interaction point. Also the polar angle of the track with respect to the beam pipe (θ) is required to be larger than 12 mrad to reject badly reconstructed tracks. An extra cut on the impact parameter $IP_{PU}/\sigma_{IP_{PU}} > 3$ with respect to any other reconstructed primary vertex (called *Pile-Up* vertex and labeled PU) is required. All the cuts mentioned are summarized in Table 3.1. These cuts select rather loosely a sample of tagger candidates that are analyzed in a second step with the aim of optimizing the tagging performance.

Tagging particle	
Variable	selection cut
p	< 200 GeV/c
p	> 2 GeV/c
p_T	< 10 GeV/c
track $\chi^2/d.o.f$	< 5
track type	*long* or *upstream*
charge	$\pm$
θ	> 12 mrad
$\lvert\Delta\phi\rvert$	> 5 mrad
$IP_{PU}/\sigma_{IP_{PU}}$	> 3
other	not in the signal B decay chain

TABLE 3.1: Cuts applied to select any tagger candidate.

A reliable particle identification (PID) is of paramount importance for both the event selection and the flavour tagging. This is provided in LHCb by the two Ring Imaging Cherenkov (RICH) devices, the Calorimeters and Muon system. Likelihoods for the kaon, pion, muon and electron hypotheses are calculated from the the information of the relevant detectors. The differences between the logarithm of the likelihood for the muon (or electron, kaon, proton) and the pion hypotheses (referred as $\Delta LL_{\mu-\pi}$, $\Delta LL_{e-\pi}$, $\Delta LL_{k-\pi}$ and $\Delta LL_{p-\pi}$ in the following) are used to refine the selection of the tagger candidates.

An iterative process can be applied over the preselected tagging particles to select the best possible tagging candidate. This process allows us to select a set of cuts which maximize the tagging performance for a given tagger by counting the number of right and wrong tagged events for different cuts on different variables. In the following sections, the opposite side and same side taggers are explained in detail. The internal structure of the taggers and the cuts which maximize the performances on MC10 are shown for $B^+ \to J/\psi K^+$ signal events.

[2]Requires hits in the Velo and TT-stations. Only used in the secondary vertex charge tagger.

3.3 Opposite side flavour tagging

Opposite side tagging algorithms are used to tag the flavour of the reconstructed B meson by looking at the flavour of the opposite b hadron. Two types are used: single particle taggers, when the charge of the lepton from semileptonic B_{opp} decays, or the charge of a kaon from $b \to c \to s$ decay chain are used; and the inclusive reconstruction of secondary vertices, when the charge of the inclusive secondary vertex reconstructed from B_{opp} decay products (also referred as "vertex charge" tagger) is used. When the accompanying b hadron is a neutral B meson, all these methods have an intrinsic dilution due to the possibility of flavour oscillations.

In Table 3.2 the B content from the opposite side is shown. Notice that around in 50% of the events a neutral B meson is produced, so an intrinsic dilution due to flavour oscillations can occur.

B Opposite side meson composition (%)	
B_d 39.1 ± 0.7	B^+ 39.1 ± 0.7
B_s^0 9.58 ± 0.5	B_c^+ 0.12 ± 0.06
Baryons 12.1 ± 0.6	

TABLE 3.2: Distribution of B_{opp} types from MC simulations.

The opposite B meson is nearly independent of the signal B, so the opposite side content should be channel-independent, however trigger conditions could affect the B_{opp} momentum and could bias the sample.

The muon, electron, kaon and vertex charge taggers are the 4 opposite side taggers used by the *FT* package at LHCb and are detailed in the following sections. At the end of this Chapter, we summarize some other opposite side taggers that had been developed but are not used due to their low effective efficiency or because they are correlated with other existing taggers. This was done with the aim of documenting all the work done with respect to the flavour tagging algorithms and with the hope to improve the algorithms in the future.

3.3.1 Opposite side Muon Tagger

For the opposite side muon tagger, a muon from the semileptonic B decay is used to tag the initial B meson. Assuming that no oscillation occured in the opposite B, the charge of the muon indicates the flavour of the reconstructed B meson at the decay point.

For a good particle identification, only tracks with $\Delta LL_{\mu-\pi} > 1$ are identified as muons. To reduce the number of reconstructed muons that could come from any other source, different cuts are applied: a $\chi^2/d.o.f < 3.2$ requirement is applied to the reconstructed tracks from the tracking system and a $p_T >$ 1.2 GeV/c is required to reduce the contribution from $b \to c \to l$

decays which would tag the wrong charge. In Figure 3.3 these variables are plotted and the ϵ_{eff} computed for a given cut in the variable.

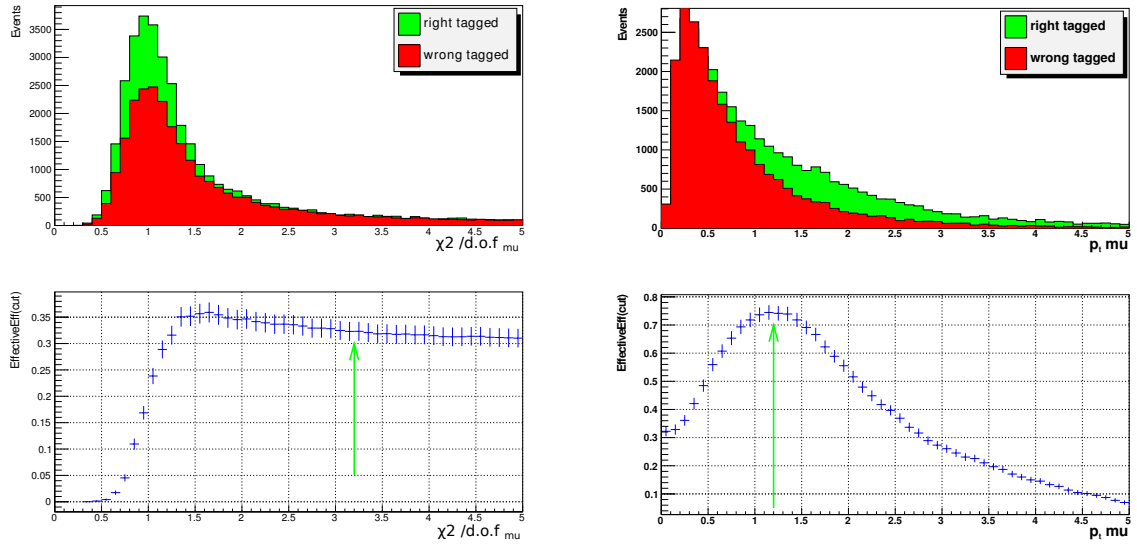


FIGURE 3.3: Tagging muon candidate $\chi^2/d.o.f$ and p_T distributions for MC10 simulated events. Above: right (green) and wrong (red) tagged events. Below: Effective efficiency for each cut on the mentioned variable. The arrow indicates the MC10 optimization cut.

An additional algorithm, called `Non Shared Hits` (NSH) [79], is used to increase the performances. It was developed in order to avoid the contamination due to fake muons (actually true pions) due to close-by tracks sharing the same hits of a true muon track, giving rise to an extra muon that can be eligible to be an opposite side tagger. This algorithm is particularly useful in all decay channels where there is a muon in the final state. Its output is used to veto the use of such tracks as muon taggers. In addition, every track in LHCb has a field called `clone distance` which represents the 'distance' to all the other tracks in the event, evaluated using the Kullback-Liebler criterion [80]. If any of the tracks is flagged as a clone it will have a non zero value. To avoid the selection of clone muon candidates only muons with a null clone distance are used. The use of these two algorithms shows an increase in the effective efficiency of about 0.1%.

If more than one candidate is selected, only the muon with highest p_T is used as tagger. All the cuts applied in the opposite side muon tagger are sumarized in Table 3.3.

3.3.2 Opposite side Electron Tagger

Analogously to the muon tagger, in the opposite side electron tagger, the electron from the semileptonic B opposite decay is used to tag the initial B meson. The charge of the electron indicates the flavour of the reconstructed B meson.

OS muon selection cuts
$\Delta LL_{\mu-\pi} > 1$
$p_T >$ 1.2 GeV/c
track $\chi^2/d.o.f < 3.2$
only *long* tracks
NSH and no clones

TABLE 3.3: Selection cuts for OS muon.

To maximize the tagging power, only tagging particles with a $\Delta LL_{e-\pi} > 4$ and in the HCAL acceptance are identified as electrons. In addition, only particles with transverse momentum $p_T >$ 1.2 GeV/c are selected (see Figure 3.4).

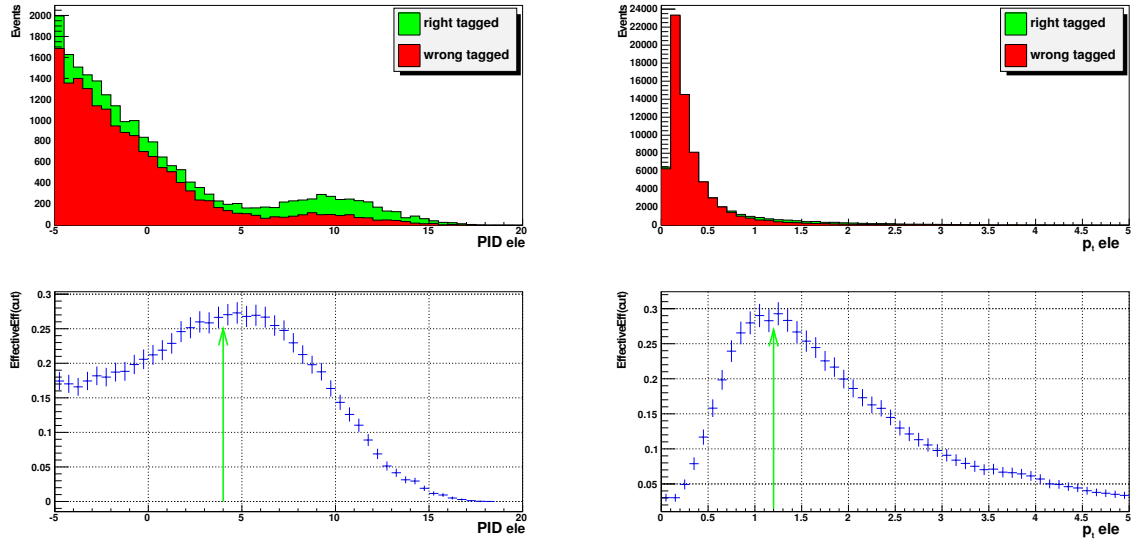


FIGURE 3.4: Tagging electron candidate $\Delta LL_{e-\pi}$ and p_T distributions for MC10 simulated events. Above: right (green) and wrong (red) tagged events. Below: Effective efficiency for each cut on the mentioned variable.

Moreover, a maximum ionization charge deposited in the silicon layers of the Vertex Locator (VELO) is required. This helps reducing the background components coming from photon conversions close to the interaction point. For the electron identification a useful variable is the ratio of the particle energy E, measured in the electromagnetic calorimeter, and the momentum p of the candidate electron, measured with the tracking system. A cut on $E/p > 0.8$ is applied to optimize the selection of the tagger candidates. These two variables are shown in Figure 3.5.

Again, if more than one candidate is selected, only the electron with highest p_T is used as tagger. The cuts for the opposite side electron tagger are sumarized in Table 3.4.

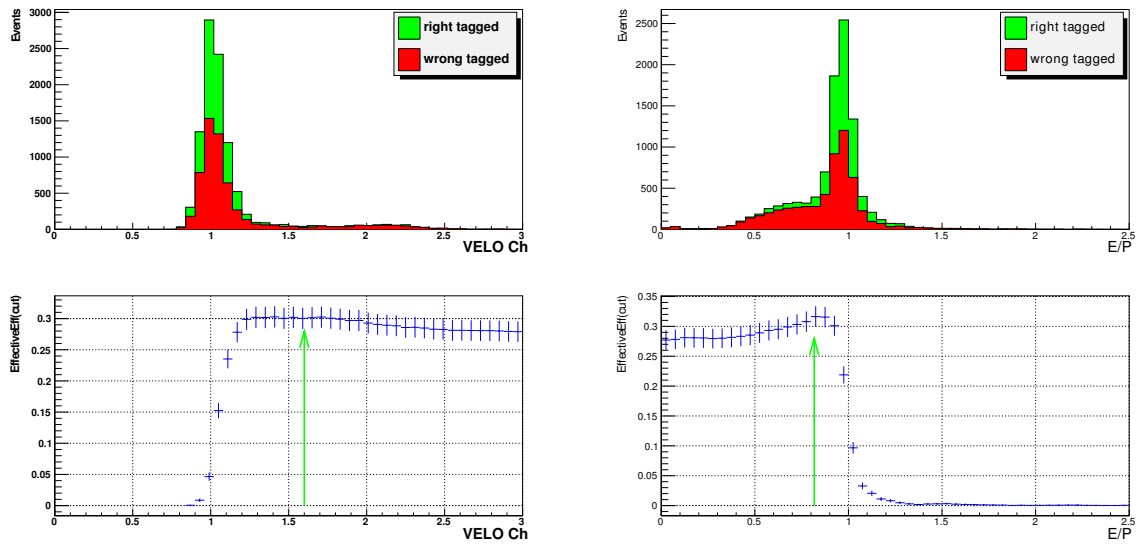


FIGURE 3.5: Tagging electron candidate VELO charge and E/P distributions for MC10 simulated events. Above: right (green) and wrong (red) tagged events. Below: Effective efficiency for each cut on the mentioned variable.

OS electron selection cuts
$\Delta LL_{e-\pi} > 4$
$p_T >$ 1.2 GeV/c
$E/p >$ 0.8
$VELO_{charge} < 1.6$
only if in Hcal Acceptance
only *long* tracks

TABLE 3.4: Selection cuts for OS electron.

3.3.3 Opposite side Kaon Tagger

In case of the opposite side kaon tagger, the kaon from the $b \to c \to s$ decay chain is used. The charge of the kaon is the opposite of the B_{opp} meson and so the same as the reconstructed B assuming no oscillation on the B_{opp} meson.

For the kaon candidate it is required a transverse momentum $p_T > 0.5$ GeV/c, an impact parameter with respect to the primary vertex in absolute value $|IP| < 1.25$ mm, an impact parameter with respect to the primary vertex with significance $IP/\sigma_{IP} > 3.35$ and a tighter cut on the track $\chi^2/d.o.f < 2.75$. These cuts enhance the contribution of kaons from b meson decays with respect to the kaons produced in the fragmentation and maximize the ϵ_{eff} as shown in Figure 3.6.

To ensure a good identification of kaon versus pions and protons, only particles with $\Delta LL_{k-\pi} > 6$ and $(\Delta LL_{k-\pi} - \Delta LL_{p-\pi}) > -4$ are identified as kaons. The Kullback-Liebler criterion [80] is

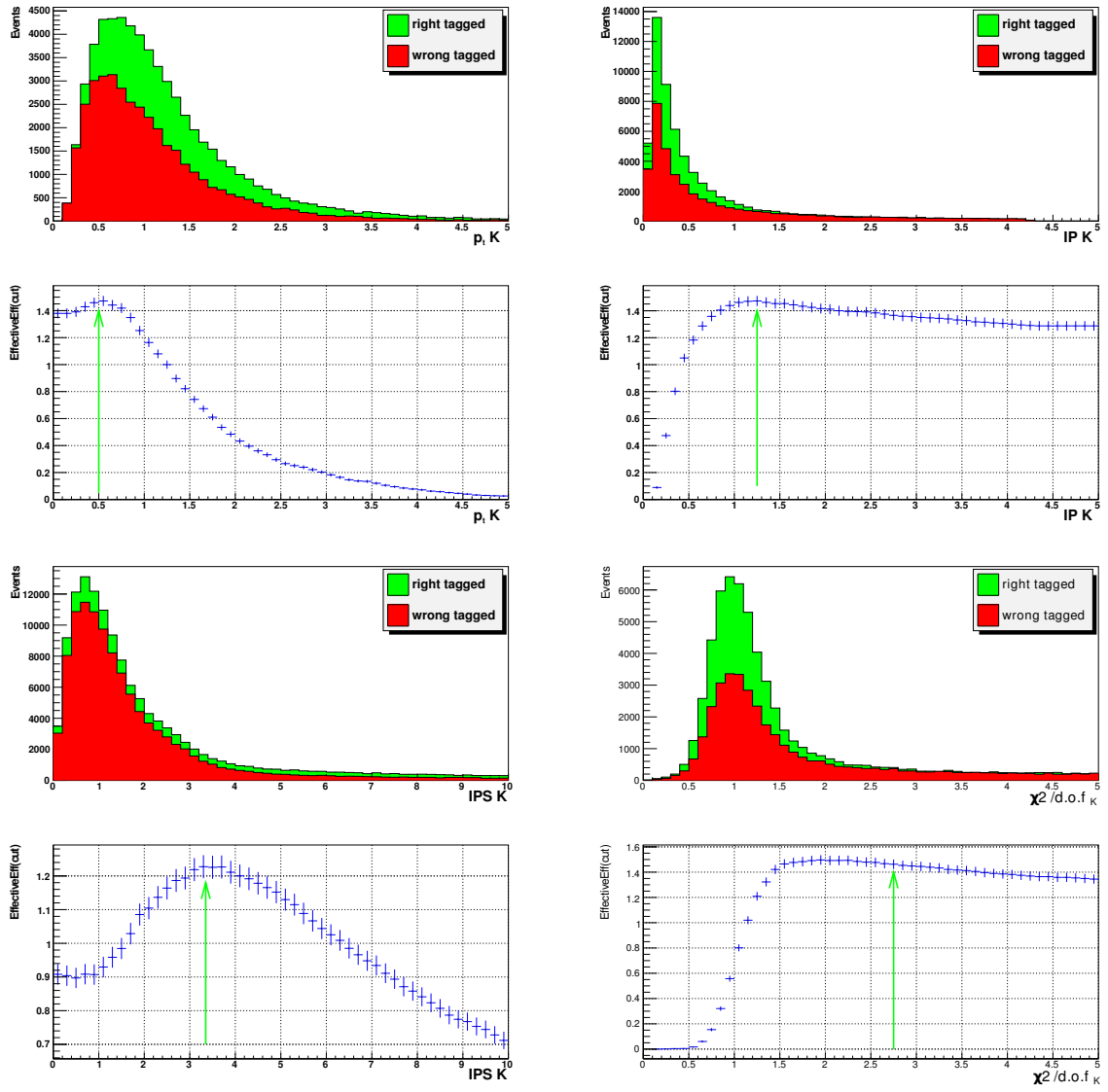


FIGURE 3.6: From top left to bottom right: kaon p_T, $|IP|$, IP/σ_{IP} and $\chi^2/d.o.f$ distributions for MC10 simulated events. Above: right (green) and wrong (red) tagged events. Below: Effective efficiency for each cut on the mentioned variable.

also used to exclude clone particles. Finally an extra cut on the impact parameter with respect to any other reconstructed primary vertex (Pile-Up) with significance $IP_{PU}/\sigma_{IP_{PU}} > 4.5$ has been applied.

If more than one candidate is selected, only the one with highest p_T is used. An extra cut on the estimated probability of the tagger to be correct (detailed in 3.7) is applied at $Prob^{kaon}_{min} > 0.54$. The cuts applied in the opposite side kaon tagger are sumarized in Table 3.5.

OS kaon selection cuts
$\Delta LL_{k-\pi} > 6 \quad (\Delta LL_{k-\pi} - \Delta LL_{p-\pi}) > -4$
$p_T > 0.5$ GeV/c
track $\chi^2/d.o.f < 2.75$
only *long* tracks no clones
$IP/\sigma_{IP} > 3.35 \quad \lvert IP \rvert < 1.25$ mm
$IP_{PU}/\sigma_{IP_{PU}} > 4.5$
$Prob^{kaon}_{min} > 0.54$

TABLE 3.5: Selection cuts for OS kaon.

3.3.4 Opposite side Vertex Charge Tagger

The opposite side secondary vertex charge tagger is based on the inclusive reconstruction of a secondary vertex corresponding to the opposite B hadron decay. Only particles passing the loose preselection defined in 3.1 can be part of the reconstructed secondary vertex. The process goes as follows:

- a 2-track seed is build for the secondary vertex (SV) from all possible track candidates,
- after some geometric and kinematical cuts, more tracks are added to the seed
- the weighted charge of the reconstructed secondary vertex is computed to obtain a tagging decision
- different set of cuts on the reconstructed SV are applied to improve the effective efficiency (ϵ_{eff}) of the tagger

The reconstruction of the secondary vertex starts using two tracks as initial seed. These tracks must satisfy kinematic and geometrical cuts which enhance the probability to come from the opposite B hadron decay. These cuts were designed on MC so that we can easily compute the probability of the track to come from the opposite B by using the generator information of the associated MC particle [81].

The cuts which maximize the probability of the track to come from the opposite B: one of the two tracks must be a long track with an associated $\chi^2/d.o.f < 2.5$; the error on the impact parameter of the tracks with respect to the primary vertex (labeled as PV) must be $\sigma_{IP_{PV}} < 1$ to discard badly reconstructed tracks; the $IP_{PV}/\sigma_{IP_{PV}}$ is required to be in the range $(2.5, 100)$ to consider only tracks which are not coming from the PV; the track must have a $IP_{PV} < 3$ mm and the transverse momentum be greater than $p_T > 0.15$ GeV/c; moreover, the two tracks must be separated in ϕ more than 1 mrad and at least one track must have a $p_T > 0.3$ GeV/c.

Track pairs are then combined and fitted to a vertex. If the fit succeeds and has a $\chi^2/d.o.f < 10$, the pair will be considered as a seed to form the inclusive secondary vertex. The seed is required

to point to the forward direction with respect to the PV ($z > 0$), to be in the detector acceptance (10 mrad $< \theta <$ 350 mrad) and it is normalized by the variance of the z distribution with respect to the primary vertex (r-distance). This maximizes the probability of the seed to come from the b hadron. The mass of the seed must be greater than 0.2 GeV/c^2, while track pairs compatible with a K^0_S decay are excluded.

A likelihood for all the possible seeds is built based on the $\chi^2/d.o.f$ of the vertex, the minimum p_T, the maximum IP_{PV} and the minimum $IP_{PV}/\sigma_{IP_{PV}}$ of both tracks, as well as the $\Delta\phi$ of the two tracks, the r-distance, and the θ direction of the seed with respect to the PV. Finally, the 2-track seed with the maximum likelihood is taken if this likelihood is greater than 0.6 (see Figure 3.7), enhancing the probability of the seed to come from the opposite b hadron. At this point, the *MultiLayer Perceptron (MLP)* [82] and the *Toolkit for Multivariate Data Analysis (TMVA)* [83] packages were used to obtain a better method to distinguish between seeds coming or not from the opposite b hadron. However, the simpler likelihood method turned out to perform equally well.

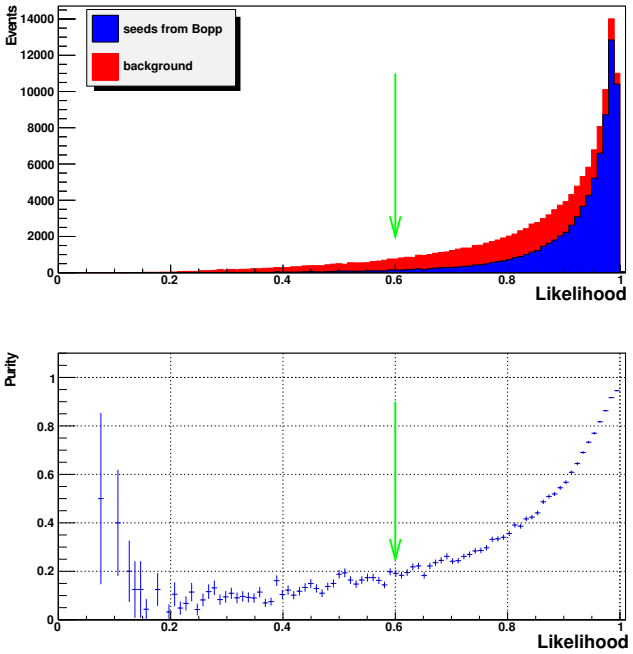


FIGURE 3.7: Above: Reconstructed seeds coming (blue) or not coming (red) from the opposite B. Below: Fraction of seeds coming from the opposite b hadron. The (green) arrow shows the cut applied in the likelihood value.

The kinematic cuts applied to form a 2-track seed are summarized in Table 3.6. In 48% of the events a seed is found. The probability of the seed to be formed by two tracks both coming from b hadron decay products is 62%.

Secondary vertex Seed	
Variable	seed request
$\sigma_{IP_{PV}}$	< 1
$IP/\sigma_{IP_{PV}}$	$2.5 < IP/\sigma_{IP_{PV}} < 100$
p_T	$p_{Tmin} > 0.15$ GeV/c $\quad p_{Tmax} > 0.3$ GeV/c
IP_{PV}	$< 3mm$
long track $\chi^2/d.o.f$	< 2.5
upstream track $\chi^2/d.o.f$	< 5
$\lvert\Delta\phi\rvert$ of 2 tracks	> 1 mrad
fit $\chi^2/d.o.f$	< 10
r-distance	$0.3 < rdist < 10$
seed mass	> 0.2 GeV
others	at least one *long* track, no duplicates, seed situated after the PV and in the acceptance, no clones, K^0_S veto.
likelihood	> 0.6

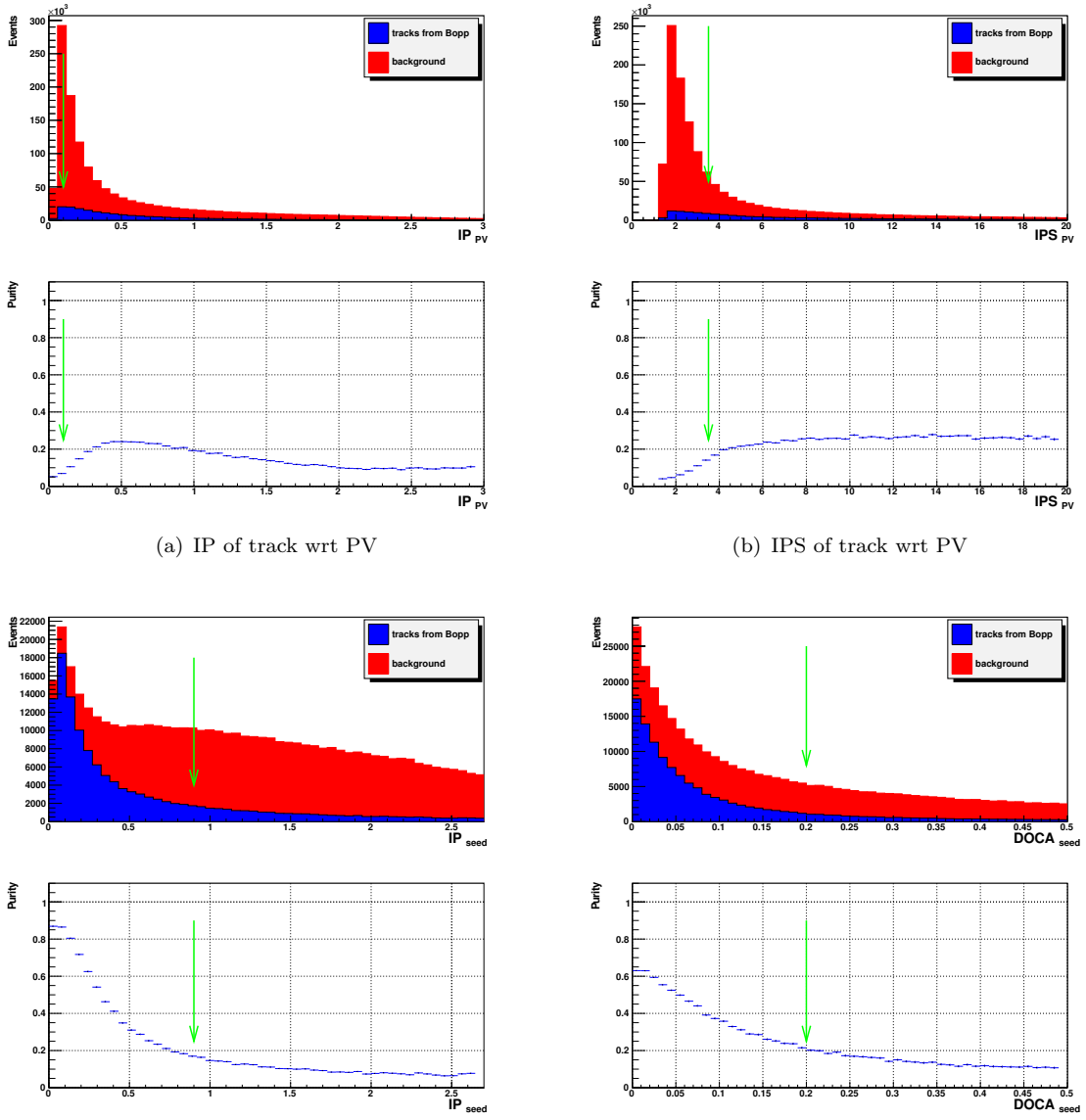
TABLE 3.6: Different set of cuts applied to the 2-track seed to select the most probable seed coming from the opposite b hadron.

If a seed is formed, other tracks can be included in the secondary vertex if they satisfy additional geometric and kinematic criteria. All the 'tagging particles' are considered, but they are only added to the vertex if the track is not a clone, the track $\chi^2/d.o.f < 3$, the $IP_{PV} > 0.1$ mm and the $IP_{PV}/\sigma_{IP_{PV}} > 3.5$. Also, more geometric criteria are requested with respect to the seed: $IP_{seed} < 0.9$ mm and the minimum distance of closed approach of the track with respect to any track in the seed has to be $DOCA_{seed} < 0.2$ mm. The distribution of some of these variables can be seen in Figure 3.8.

All the tracks that pass the mentioned above criteria (summarized in Table 3.7) are added to the inclusive secondary vertex. The probability of an extra track added to the seed to be a true b hadron decay product from the opposite b hadron is 62%. At this point two other methods were also tried: refit the 2-track seed with the extra track and apply later a $\chi^2/d.o.f$ cut or use a neural network to decide whether to add a track or not. We found no improvement in the tagging performance with these two methods, so we kept the simplest one.

Adding tracks to the seed	
Variable	selection cut
$IP/\sigma_{IP_{PV}}$	> 3.5
IP_{PV}	> 0.1
track $\chi^2/d.o.f$	< 3.0
IP_{seed}	< 0.9 mm
$DOCA_{seed}$	< 0.2 mm
others	no clones

TABLE 3.7: Cuts applied to any extra track to be added to the 2-track seed to form an inclusive secondary vertex.



(a) IP of track wrt PV

(b) IPS of track wrt PV

(c) IP of track wrt Seed

(d) DOCA of track wrt Seed

FIGURE 3.8: Distributions and cuts applied to the tracks added to the seed. The distributions and the cuts on $IP_{PV} > 0.1$ mm and $IP_{PV}/\sigma_{IP_{PV}} > 3.5$ are shown in (a) and (b). Also, cuts on track variables with respect to the seed $IP_{seed} < 0.9$ mm and $DOCA_{seed} < 0.2$ mm are shown in (c) and (d). For each Figure tracks coming from an opposite b hadron (blue) and background tracks (red) are shown (first row), below it can be seen the purity of tracks coming from opposite b hadrons for a given value of the mentioned variables.

We also tried to improve this estimator by adding tracks from fragmentation in a cone around the opposite b hadron flight direction. But these additional tracks have the effect of diluting the information carried by the tracks associated to the inclusive vertex, and no clear improvement is seen.

Weighted charge and tag decision

The inclusive reconstruction of the accompanying B decay vertex is used to determine the B hadron charge. The weighted vertex charge is defined as the normalized sum of the charges of all its tracks weighted with ${p_T}^k$:

$$Q_{vtx} = \frac{\Sigma_i p_T^k(i) Q_i}{\Sigma_i p_T^k(i)} \tag{3.11}$$

The vertex charge defined by Eq. 3.11 is used to define the tagger decision d. The k parameter is optimized together with a cut on the non-discriminating central part of the Q_{vtx} distribution, in order to maximize the effective tagging efficiency. The optimization was firstly performed over 500 thousand MC $B^+ \to J/\psi K^+$events and can be seen in Figure 3.9. This leads to $k = 0.4$ and to consider events with $|Q_{vtx}| < 0.25$ as untagged, reducing the tagging efficiency (ϵ_{tag}) but increasing the ϵ_{eff}. Other possibilities to weight the vertex charge were envisaged in the past [85] but none gave a better performance.

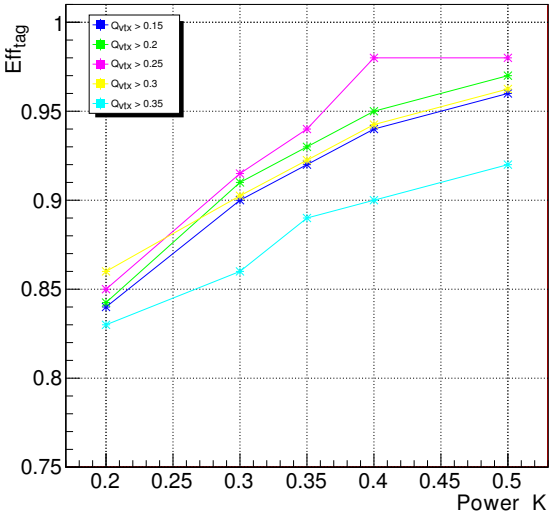
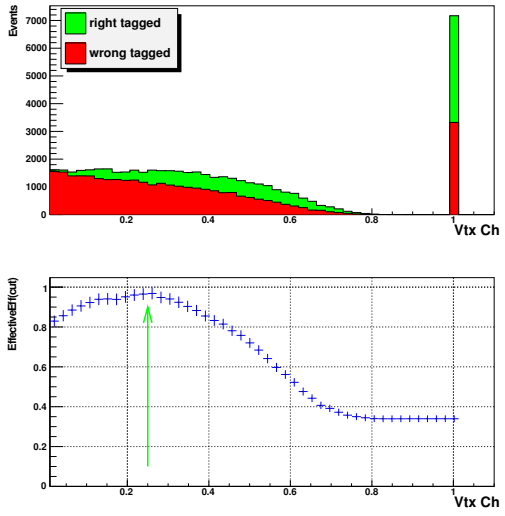


FIGURE 3.9: Effective tagging efficiency for different values of k and Q_{vtx} for MC $B^+ \to J/\psi K^+$events.

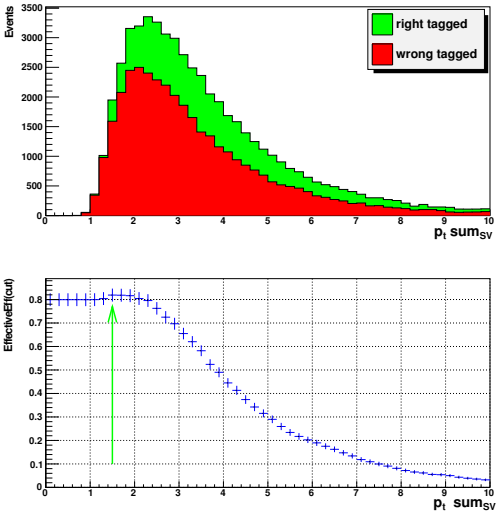
Figure 3.10(a) shows the distribution in absolute value of the weighted vertex charge when the decaying b hadron is charged. The peak at $Q_{vtx} = 1$ is populated by events with the vertices formed by tracks of the same charge. The cut on $|Q_{vtx} > 0.25|$ is chosen as the value which maximizes the effective efficiency of the vertex charge tagger given the weight $k = 0.4$.

Once the secondary vertex is reconstructed we can see that there are variables which show a dependence between right and wrong tags, so a cut on these variables might increase the effective tagging efficiency. Different variables were envisaged, a cut on the momentum and the mass of

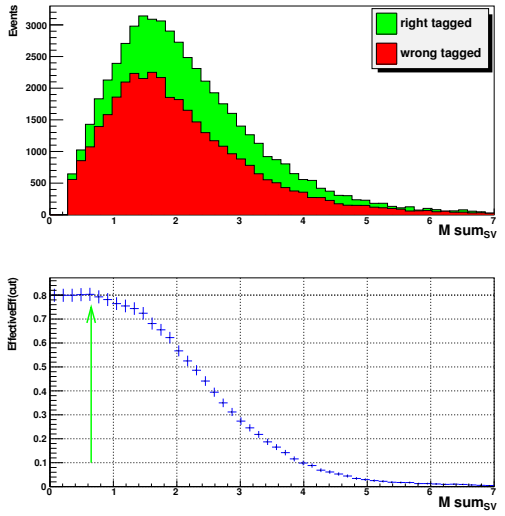
all the tracks in the secondary vertex, $P_{SV} > 10$ GeV/c and $M_{SV} > 0.5$ GeV/c^2 respectively maximizes the performances. The difference between the ϕ and θ angle of the opposite b hadron (obtained from the position of the reconstructed PV and the 2-track seed) and the corresponding of the resulting sum of all the tracks in the SV did not give a clear improvement and were finally not used. Three more variables were used instead: the sum of p_T, $IP_{PV}/\sigma_{IP_{PV}}$ and $DOCA_{seed}$ of all the tracks in the secondary vertex, leading to consider only reconstructed secondary vertices with $p_{T\,sumSV} > 10$ GeV/c, $IP_{sumSV}/\sigma_{IP_{sumSV}} > 10$ and $DOCA_{sumSV} < 0.5$ mm. Some of these variables are shown in Figure 3.10.



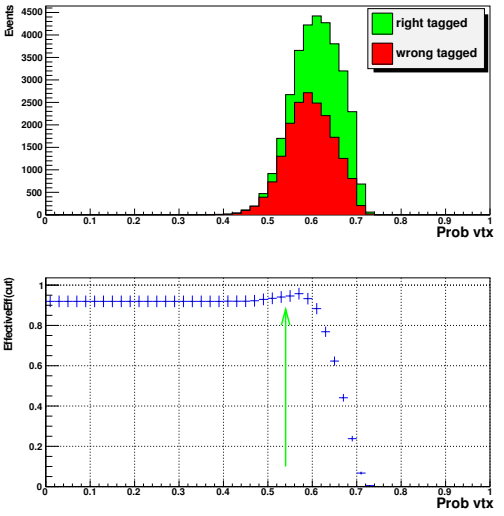
(a) Vertex charge of the inclusive reconstructed secondary vertex



(b) p_T of all tracks in SV



(c) M of the reconstructed quadrimomentum of the SV



(d) Probability of a right tag for vertex charge tagger

FIGURE 3.10: Above: distribution for right (green) and wrong (red) tagged MC10 simulated events. Below: Effective efficiency as a function of the cut applied.

To maximize the effective efficiency of the tagger an extra cut on the probability of the tagger to

be correct ($Prob^{vtx}_{min} > 0.54$) is applied. A detailed description of how this probability ($Prob^{vtx}_{min}$) is obtained is given in Section 3.7. The set of cuts applied to the inclusive reconstructed secondary vertex are summarized in Table 3.8.

Inclusive reconstructed secondary vertex cuts	
Variable	selection cut
$\|Q_{vtx}(k)\|$	$> 0.25 \quad (k = 0.4)$
P_{SV}	> 10 GeV/c
M_{SV}	> 0.5 GeV
$p_{T\,sumSV}$	> 10 GeV/c
$IP_{sumSV}/\sigma_{IP_{sumSV}}$	> 10
$DOCA_{sumSV}$	< 0.5
$Prob^{vtx}_{min}$	> 0.54

TABLE 3.8: Cuts applied to inclusive reconstructed secondary vertex, to consider it as a tagger.

3.4 Same side flavour tagging

Same side tagging algorithms determine directly the flavour of the signal B meson exploiting the correlation in the fragmentation chain. In the case of a B^0 ($\bar{b}d$), from the fragmentation of a $\bar{b}$ quark, an extra $\bar{d}$ is available to form a pion. The pions can also be originated from decays of excited B states to the signal B^0 like $B^{**+} \to B^{(*)0}\pi^+(n\pi)$ and $B^{**0} \to B^{(*)+}\pi^-(n\pi)$. These pions are used as a tagger in the so called "same side pion tagger" (SSπ). In the case of B^0_s mesons ($\bar{b}s$), an extra $\bar{s}$ is available to form a hadron which leads to a charged kaon in about 50% of the cases, the so called "same side kaon tagger" (SSK). Both hadronization proccesses can be seen in Figure 3.11. In case of B^+ (B^-) mesons and extra $\bar{u}$ (u) quark is available to form a π^- (π^+) or a K^- (K^+).

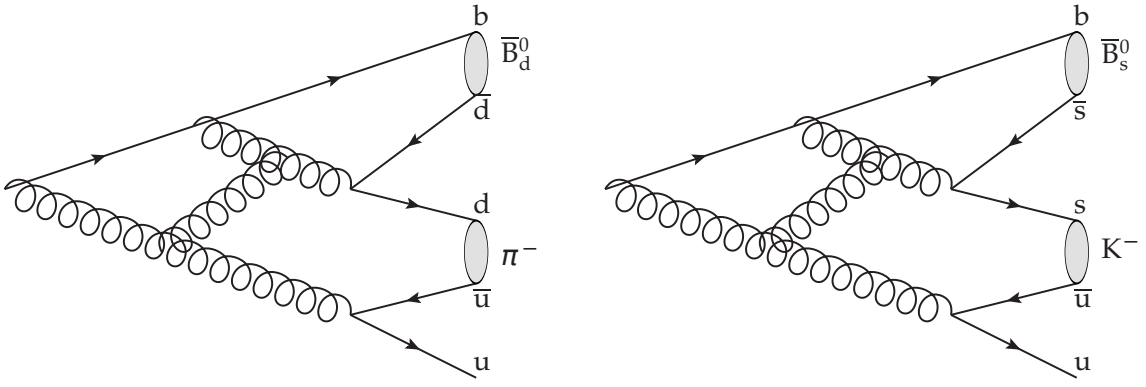


FIGURE 3.11: Feynman diagram of the B^0 and B^0_s hadronization. For B^0 mesons a pion can be originated from the extra d quark. The SSπ can be also originated from excited B states. For B^0_s mesons an extra s quark is available to form a hadron, which leads to a charged kaon in about 50% of the cases.

3.4.1 Same side Pion tagger

In B^0 events a π can be originated in the fragmentation chain from hadronization processes were an extra d quark is available (see Figure 3.11 (left)), but also from excited states of the B^0. The relative fractions of weakly decaying B mesons results in 21.0%, 62.9% and 16.2% for B^0, B^{0*} and B^{0**} respectively [62]. Of all pions coming from the B^{0**} decay, 69% are in the LHCb acceptance but only 19% are selected.

Same side taggers are selected with similar criteria to the ones chosen for the opposite side taggers. For the SSπ, only long tracks are considered. In addition, only identified pions with $\Delta LL_{k-\pi} < 4.5$ and $\Delta LL_{p-\pi} < 15$ are used. This ensures a good identification of pions versus kaons and protons. Candidates are required to have a $p_T >$ 0.5 GeV/c, $p >$ 2.5 GeV/c and $IP/\sigma_{IP} < 4$. The difference between the ϕ angle and pseudorapidity (η) of the track to the B_{sig} was also considered but the performance could not be improved. A cut on the difference ΔQ between the mass of the $B^0\pi$ combination and the mass of the reconstructed B^0 is required to be $\Delta Q < 1.5$ GeV/c^2. These last two variables are shown in Figure 3.12.

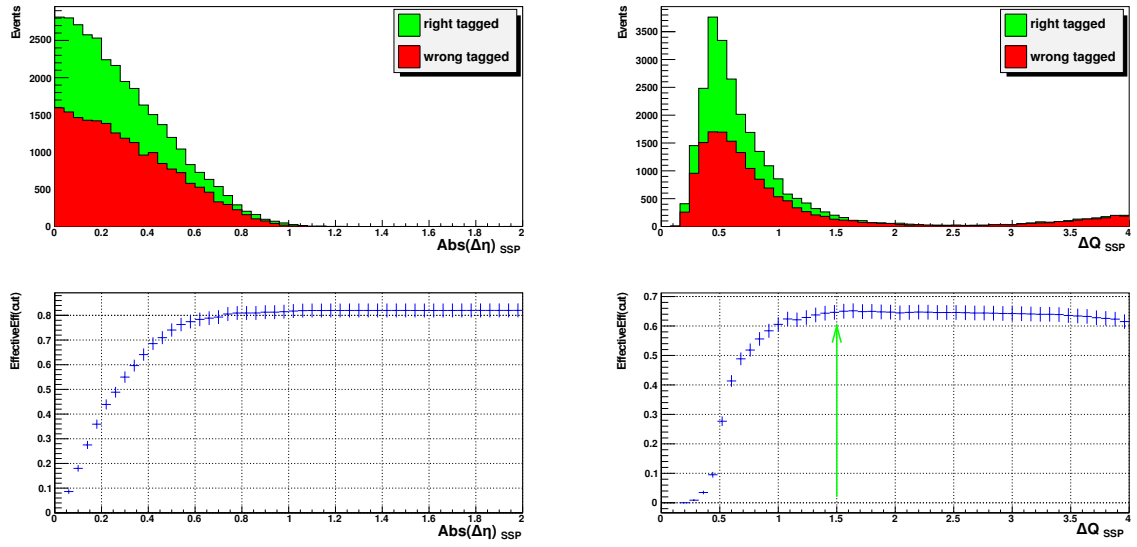


FIGURE 3.12: Distribution for MC10 simulated events. Left: difference between η of the B_{sig} and the $SS\pi$. Right: difference between the mass of the $B^0\pi$ combination and the mass of the reconstructed B^0 (ΔQ). Above: right (green) and wrong (red) tagged events. Below: Effective efficiency for each cut on the mentioned variable.

If more than one candidate is selected, only the one with highest p_T is used. Finally, an extra cut on the probability of the tagger to be correct $Prob^{SS\pi}_{min} > 0.54$ is used. The cuts are summarized in Table 3.9.

SS pion selection cuts
$\Delta LL_{k-\pi} < 4.5$ $\Delta LL_{p-\pi} < 15$
$p_T > 0.5$ GeV/c $p > 2.5$ GeV/c
only *long* tracks
$IP/\sigma_{IP} < 4.$
$\Delta Q < 1.5$ GeV/c^2
$Prob_{min}^{SS\pi} > 0.54$

TABLE 3.9: Selection cuts for SSπ.

3.4.2 Same side Kaon tagger

The same side kaon tagging algorithm determines the flavour of the signal B_s^0 meson by exploiting the correlation in the fragmentation chain. If a B_s^0 $(\bar{b}s)$ is produced in the fragmentation of a $\bar{b}$ quark, and extra $\bar{s}$ is available to form a hadron, which leads to a charged kaon in about 50% of the cases. Even if a K^{*0} is produced, it leads to a right-sign K^+. However, if a ϕ is produced, the probability to have a correct tag is 50%. In Figure 3.11 (right) the fragmentation process of a B_s^0 meson is shown.

These kaons are selected requiring a momentum $p > 5.25$ GeV/c, $p_T > 0.75$ GeV/c, an impact parameter with respect to the primary vertex with significance $IP/\sigma_{IP} < 4.125$ and a tighter cut on the track $\chi^2/d.o.f < 3.75$. These cuts enhance the contribution of kaons produced in the fragmentation. For the same side kaon tagger cuts on new variables with respect to the signal B are introduced (see Figure 3.13): a difference in the pseudo-rapidity with respect to the reconstructed B signal $|\Delta\eta| < 0.525$; a difference in the ϕ angle $|\Delta\phi| < 0.7$; a cut on $|\Delta R| < \sqrt{\Delta\eta^2 + \Delta\phi^2} < 1.2$; and finally $\Delta Q < 1.5$ GeV/c^2, where ΔQ is the difference between the mass of the $B_s^0 K$ system and the mass of the reconstructed B_s^0.

To ensure a good identification of pions versus kaons and protons, only identified kaons with $\Delta LL_{k-\pi} > 4.5$ and $(\Delta LL_{k-\pi} - \Delta LL_{p-\pi}) > -8.5$ are used. If more than one candidate is selected, only the one with highest p_T is used. The cuts were optimized with the $B_s^0 \to D_s^- \pi^+$ channel and are summarized in Table 3.10.

SS kaon selection cuts
$\Delta LL_{k-\pi} > 4.5$ $(\Delta LL_{k-\pi} - \Delta LL_{p-\pi}) > -8.5$
$p_T > 0.75$ GeV/c $p > 5.25$ GeV/c
long tracks
track $\chi^2/d.o.f < 3.75$
$IP/\sigma_{IP} < 4.125$
$\Delta\eta < 0.525$ $\Delta\phi < 0.7$ $\Delta R < 1.2$
$\Delta Q < 1.5$ GeV/c^2

TABLE 3.10: Selection cuts for SSK.

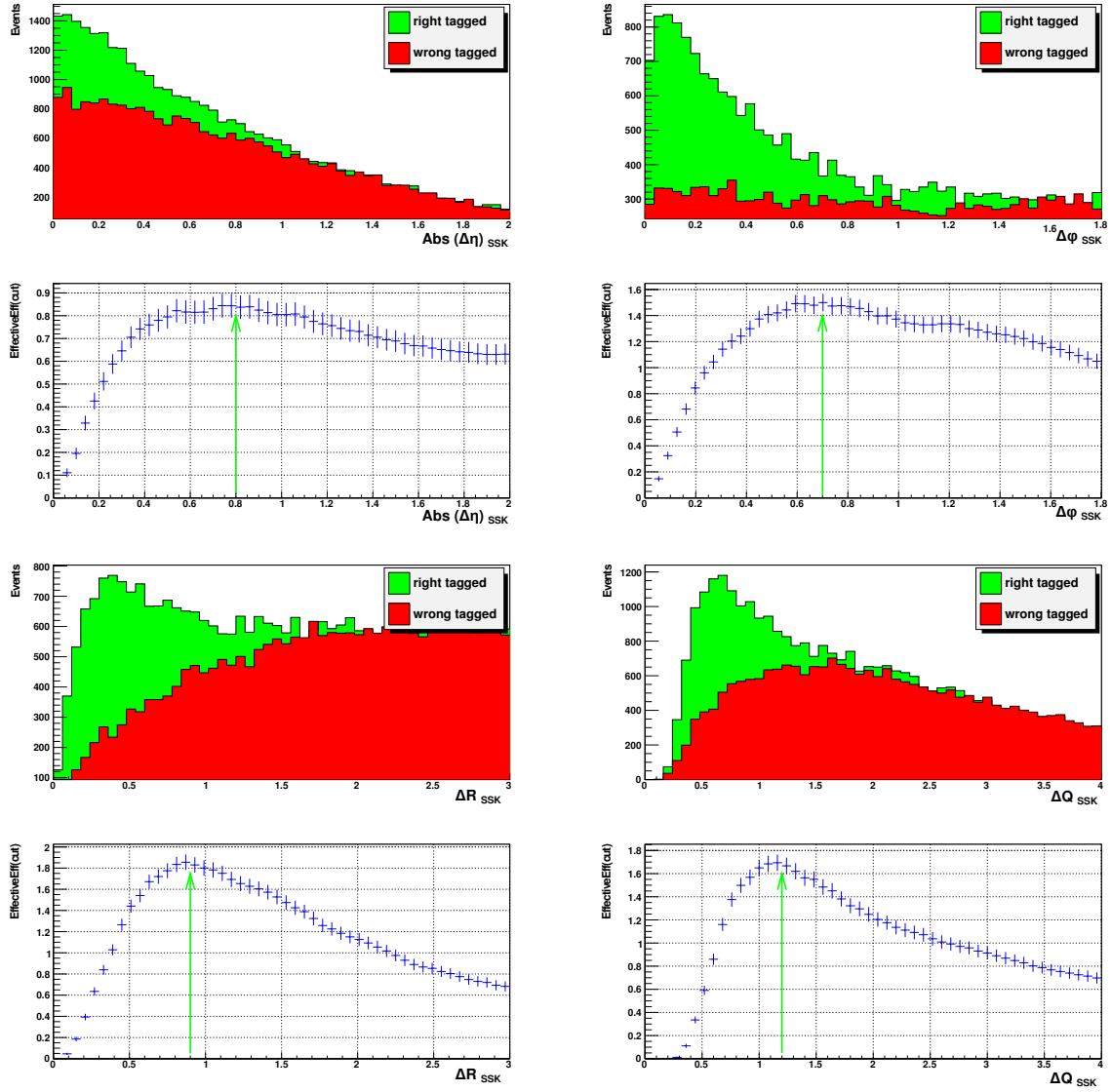


FIGURE 3.13: Distribution for MC10 simulated events. From top left to bottom right: Difference between η angle, ϕ angle, ΔR of the B_{sig} and the SSK and difference between the mass of the B^0K combination and the mass of the reconstructed B^0 (ΔQ). Above: right (green) and wrong (red) tagged events. Below: Effecive efficiency for each cut on the mentioned variable.

3.5 Optimization of taggers for high multiplicity events

To understand how taggers behave in the case of high multiplicity events, their performances were studied for different number of primary vertices and different tagging candidate multiplicities on MC $B^+ \to J/\psi K^+$events. The proton-proton collision associated to the reconstructed B_{sig} is known as 'Primary Vertex' (PV), while the other reconstructed primary vertices are named Pile-Up vertices (PU). The ϵ_{eff} of the taggers decreases when the PU or the tagging candidate multiplicity increases. In Table 3.11 the performances for the OS taggers are shown for different number of PU vertices and tagging candidates multiplicity.

	ϵ_{tag}	ω	ϵ_{eff}
Muon tagger Performance			
All	4.93 ± 0.03	30.42 ± 0.33	0.76 ± 0.03
$PU = 1$	5.05 ± 0.07	30.08 ± 0.64	0.80 ± 0.05
$PU = 2$	4.91 ± 0.06	29.88 ± 0.55	0.79 ± 0.04
$PU \geq 3$	4.86 ± 0.06	31.19 ± 0.55	0.69 ± 0.04
$5 < Mult < 30$	4.43 ± 0.03	28.11 ± 0.44	0.85 ± 0.04
$30 < Mult < 60$	5.70 ± 0.06	33.03 ± 0.53	0.66 ± 0.04
$Mult \geq 60$	6.40 ± 0.17	35.20 ± 1.32	0.56 ± 0.10
Electron tagger Performance			
All	2.85 ± 0.03	33.65 ± 0.45	0.30 ± 0.02
$PU = 1$	2.92 ± 0.05	32.46 ± 0.86	0.36 ± 0.04
$PU = 2$	2.81 ± 0.04	33.83 ± 0.75	0.29 ± 0.03
$PU \geq 3$	2.85 ± 0.04	34.31 ± 0.73	0.28 ± 0.03
$5 < Mult < 30$	2.45 ± 0.03	32.10 ± 0.61	0.31 ± 0.02
$30 < Mult < 60$	3.40 ± 0.05	35.19 ± 0.69	0.30 ± 0.03
$Mult \geq 60$	4.17 ± 0.14	36.90 ± 1.65	0.29 ± 0.07
Kaon tagger Performance			
All	21.32 ± 0.07	36.86 ± 0.17	1.47 ± 0.04
$PU = 1$	21.65 ± 0.13	35.31 ± 0.32	1.87 ± 0.08
$PU = 2$	21.47 ± 0.11	37.03 ± 0.28	1.44 ± 0.06
$PU \geq 3$	20.96 ± 0.11	37.79 ± 0.28	1.25 ± 0.06
$5 < Mult < 30$	19.29 ± 0.08	35.43 ± 0.22	1.64 ± 0.05
$30 < Mult < 60$	25.68 ± 0.12	38.35 ± 0.26	1.39 ± 0.06
$Mult \geq 60$	20.27 ± 0.28	39.99 ± 0.76	0.81 ± 0.12
Vertex charge tagger Performance			
All	10.86 ± 0.05	35.18 ± 0.22	0.95 ± 0.03
$PU = 1$	11.29 ± 0.10	34.01 ± 0.44	1.15 ± 0.06
$PU = 2$	10.92 ± 0.08	35.63 ± 0.38	0.90 ± 0.05
$PU \geq 3$	10.50 ± 0.08	35.59 ± 0.38	0.87 ± 0.05
$5 < Mult < 30$	9.32 ± 0.06	33.69 ± 0.32	0.99 ± 0.04
$30 < Mult < 60$	13.60 ± 0.09	36.66 ± 0.35	0.97 ± 0.05
$Mult \geq 60$	12.83 ± 0.23	37.39 ± 0.95	0.82 ± 0.12

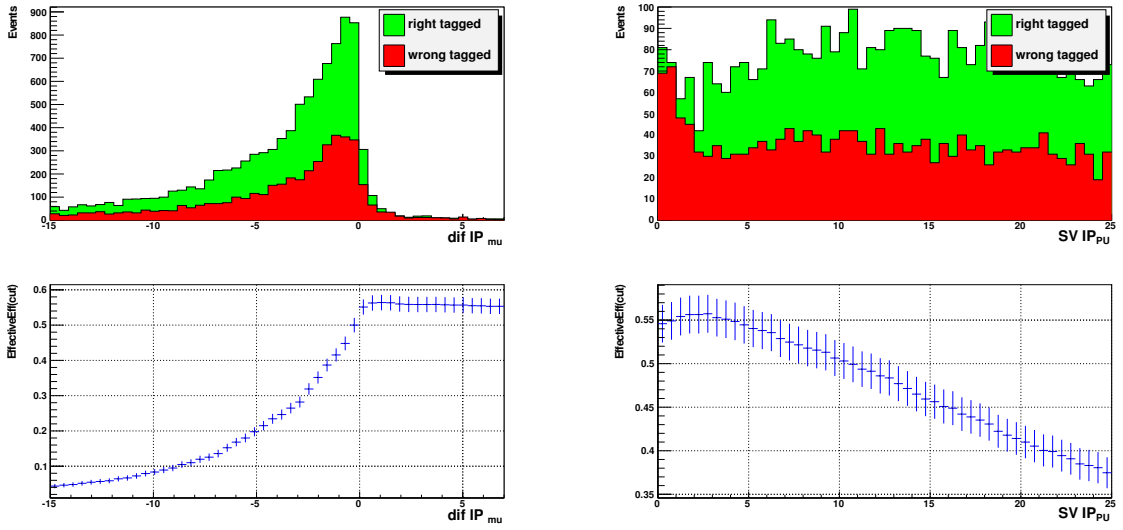
TABLE 3.11: Tagging performance for OS taggers for different number of PileUp vertices and track multiplicities on $B^+ \to J/\psi K^+$MC10 events.

This behaviour can be seen in both MC and data. The tagging performances for high multiplicity events on data was firstly studied in [86] and can be seen in Chapter 6.

In order to improve the performances for events with a high number of PU vertices, different variables (related with the extra PU vertices) are studied.

The position of the different PU vertices can be measured experimentally, and for each tagger, the impact parameter of the tagging particle with respect to these extra Pile-Up vertices was studied. In particular, the difference on the impact parameter of the tagging particle with respect to the reconstructed vertex and the PU vertex was studied. If this variable is positive it is more likely that the tagging particle comes from the PU instead of from the PV, so the probability to make a wrong tag would be higher (see Figure 3.14(a)). The same was done

considering the mean of the impact parameter with respect to all the possible PU vertices. None of these two cuts has thought a significant impact on the performance.



(a) Difference between the impact parameter with respect to the PileUp vertex and the reconstructed vertex (`dif IP`) for muon tagging candidates

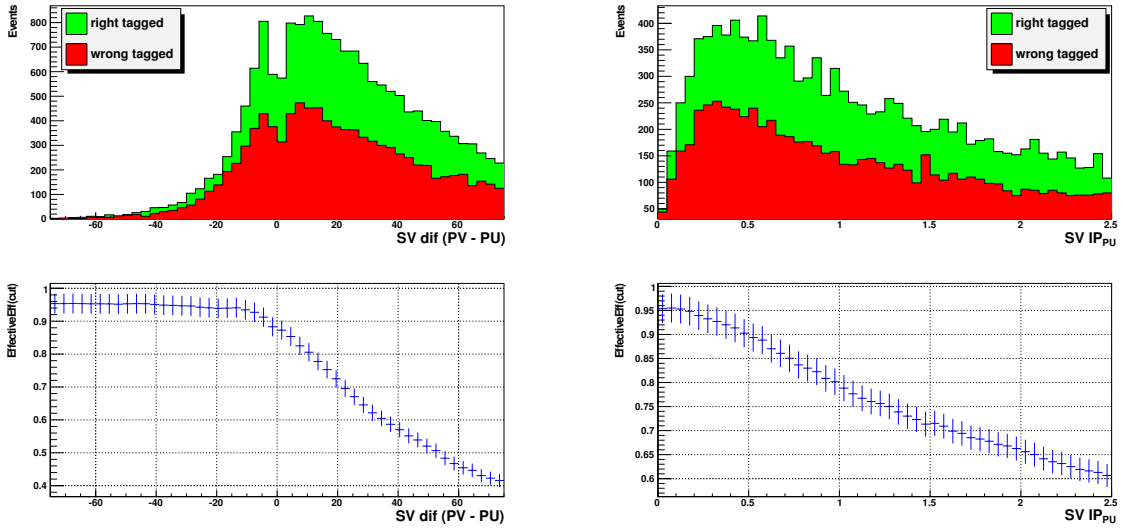
(b) Impact parameter of the reconstructed SV momentum and the closest PU vertex for muon tagging candidates

FIGURE 3.14: Above: Distribution of right (green) and wrong (red) tagged events. Below: Effective eficiency for a given cut on the variable. Here the ϵ_{eff} is related to all the events with more than 1 reconstructed primary vertex.

Another variable is the distance in the z-position between the origin of the tagging particle track and the z-position of the PU. If a track is near enough the PU, it is possible that the tagging candidate comes from this PU instead of coming from the opposite b hadron. A small improvement could be seen only for the muon tagger (Figure 3.14(b)). As the gain in ϵ_{eff} was small this variable was finally not used.

For the vertex charge tagger, different variables, which relate the secondary vertex to the PileUp vertices, were studied. In particular, we studied the inclusive Secondary Vertex (SV) position and z-position with respect to the reconstructed primary vertex and the nearest PU vertex. Also, one can *a priori* think that if the SV is nearer to the PileUp than to the reconstructed PV the inclusive reconstructed SV could be probably originated from the PU vertex instead of by the PV, causing a dilution in the tagging power. Figure 3.15(a) shows the difference in absolute value between the z-distance of the SV with respect to the PV and the distance of the SV with respect to the PU vertex. A negative value means the SV is closer to the PileUp vertex than to the reconstructed PV, and so, one could expect an improvement of the tagging power. For the same reason the IP of the SV with respect to the nearest PileUp vertex was studied as shown in Figure 3.15(b). Still, none of these variables shows a clear peak in the ϵ_{eff} graph of Figure 3.15.

Another idea was to improve the performance of the events with high PU or high track multiplicity events by introducing a different set of cuts and calibration parameters for the different



(a) Difference between the absolute value of the z-distance between (SV-PV) and (SV-PU). Negative values indicates the SV is closer to the PU than to the PV

(b) Impact parameter of the reconstructed SV with respect to the closest PU vertex

FIGURE 3.15: Above: Distribution of right (green) and wrong (red) tagged events. Below: Effective eficiency for a given cut on the variable.

PU (or tagging candidate multiplicity) conditions which would optimize the performances. It turns out that the optimal set of cuts which maximize the ϵ_{eff} are the same, independently of the number of Pile-Up vertices or track multiplicity. This check was repeated with data confirming the hypotesis.

3.6 Tagger Decision

From all the tagging candidates, each tagger selects the best tagging particles according to a different set of cuts which maximize the probability of the tagger to give the correct tagging decision. In the single particle taggers, in case of multiple candidates, the one with the highest p_T is chosen. An alternative possibility was also studied: to choose the tagging candidate according to a neural net ouput (detailed in Section 3.7.2).

For the opposite side taggers the charge[3] of the tagger identify the flavour of the opposite B, while for same side pion (or kaon) tagger it is related to the flavour of the signal meson as mentioned before. The tagging decision d is defined by the following convention: signal B hadrons containing a b quark: $d = -1$, signal B hadrons containing a $\bar{b}$ quark: $d = +1$.

A global tagging decision can be obtained by combining all the active taggers in the event. This process is detailed in Chapter 5.

[3]For the inclusive vertex charge tagger the sign of the weighted charge is considered instead.

3.7 Mistag probabilities

For each tagger i, the probability of the tag decision to be wrong or mistag probability ω_i is estimated event by event as a function of several kinematic and geometrical properties of the tagger and of the event itself. It is evaluated by means of a *neural net* trained on MC events to identify the correct flavour of the signal B meson. To correct for differences between data and MC, the probability of the tagger to be correct $p_i = 1 - \omega_i$, is calibrated on data using control channels as explained in Chapter 4. When more data will be available, the training will be performed on self tagging control channels, like $B^+ \to J/\psi K^+$. Thus, a posteriori calibration will not be needed anymore.

The probability of the tagger to be correct can be used to increase the ϵ_{eff} by selecting only events with higher p, as can be seen in the kaon, SSπ and vertex charge tagger, where a cut on $p > 0.54$ is applied maximizing the tagging power. Furthermore, this probability can be used to obtain a final tagging decision where all the taggers available in the event are combined to produce a single decision. Besides, the wrong tag estimation ω is mandatory for the tagged analysis in the CP fits.

The root *MultiLayer Perceptron (MLP)* method [82] is chosen to train the neural nets in all the taggers by looking at the correct flavour of the signal B meson (although the *Toolkit for Multivariate Data Analysis (TMVA)* package [83] was also tried giving equivalent results).

3.7.1 Neural Networks

The artificial neural nets are sophisticated and powerful tools for statistical classification and prediction. Their method is based on the biological neurons, which are composed of an input structure (dendrites), a cell body, and an output structure (the axon). If the total signal received at the cell body from the dendrites (which are linked by synapses) exceeds a certain threshold, the neuron fires an electro chemical signal along the axon. The artificial neural networks have an input vector $\vec{x}$ where their values x_i corresponds to the dendrites and are connected via weights ω_{ij} that have the same function as the synaptic efficiency. The threshold of the neuron can be modeled by a step function, a sigmoid or a fermi function (a nonlinear function enables the neural network to learn nonlinear correlations between the input variables).

Although it is possible to connect directly the input layer to the output, the neural networks used for classification and prediction have typically one additional hidden layer between the input and the ouput layers. If there are no connetions looping back to previous layers and the information is only transfered in one direction, these neural networks are referred as feed forward. This is the structure of the neurons produced by the *MLP* package used in all the taggers. The layout of the neural net used for the SS kaon tagger can be seen in Figure 3.16.

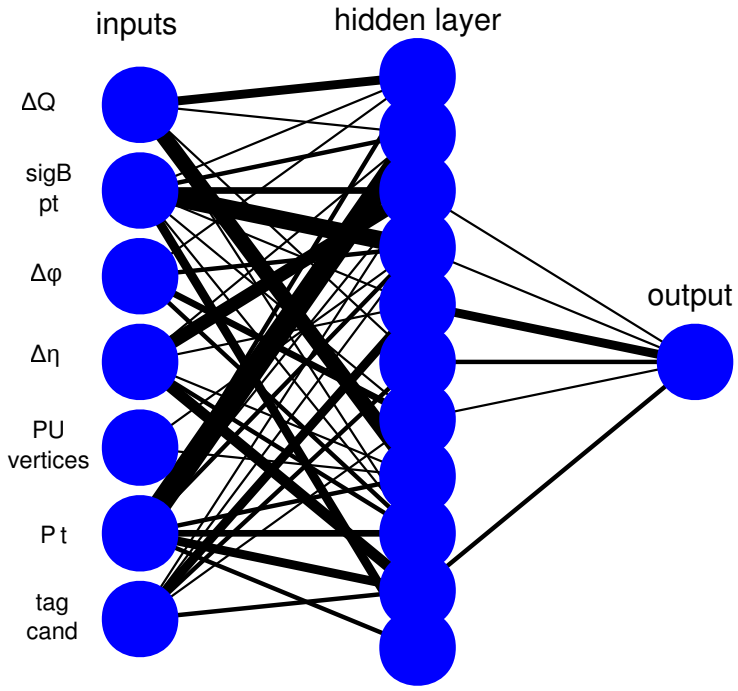


FIGURE 3.16: Layout of the artificial neural network for the SS kaon. For all the LHCb taggers the neural networks contains one input layer, one hidden layer and one output neuron.

To use a neural net for a classification problem it has to be trained first. The training process consists in finding the set of weights which maximizes the matching to the output response. For such training we need a set of data where the truth is known, called training sample. The training relates to the method of back propagation learning, in which the output is calculated for every input vector and compared to the true value. The comparison is made calculating a cost function $E(\omega_{ij})$ which describes the deviation between the neural network output $O(\omega_{ij})$ for a set of weights (ω_{ij}) and the true value (T). The most commonly used cost function is

$$E(\omega) = \Sigma_{\mu}(O_{\mu}(\omega) - T_{\mu})^2 \tag{3.12}$$

where μ is running over all training events. If T_{μ} corresponds to 1 or 0 when the decision is correct or uncorrect, the cost function has a global minimum for $O_{\mu} = p_{\mu}$ [84], where p_{μ} is the probability of the neural net to give the correct answer. In order to find its minimum up to five different methods can be used [82]. The batch mode of the *MLP* package is implemented for the neural nets of the taggers. This is the method of the steepest descent with fixed step size, which consist on a stochastic minimization where the weights are updated after considering all the examples. Thereby the change of each weight is opposite to the direction of the cost function. The weights are updated after each example according to the formula: $w_{ij}(\mu+1) = w_{ij}(\mu) + \Delta w_{ij}(\mu)$ where

$$\Delta\omega_{ij}(\mu) = -\xi\frac{\partial E}{\partial\omega_{ij}} \tag{3.13}$$

with ξ being the learning rate, which determine how fast the weights are changed.

3.7.2 Mistag estimation in the taggers

The neural nets used in the taggers to estimate the wrong tag fraction ω were not only designed to keep to a minimum the number of useful variables, but also at the same time, to optimize the discrimination between right and wrong tagged events, which can be used to compute a probability of the tagger to be correct. To choose which are the important variables, it is important to know which are the variables which show a dependency on ω for different values of the variable itself. The higher the variation in ω as a function of the variable the more useful will be the variable for the neural network. On the contrary, a flat ω distribution along a variable would indicate that the variable does not play an important role in discriminating wrong from right tags, so that it should not be used.

An iterative process has been applied: for different input variables, the weights were generated and the network efficiency studied, keeping at the end the network structure with the best efficiency and least number of variables. The neural nets input depend on event properties like the momentum of the signal B, the number of Pile-Up vertices and the preselected tagging candidates. Moreover they depend on geometrical and kinematical properties of the tagging particle or of the secondary vertex.

- For the **single particle opposite-side taggers**, like muons, electrons and kaons, the p_T of the signal B and the number of Pile-Up vertices are included as event properties, plus the particle p, p_T and its impact parameter with respect to the primary vertex. In case of the OS kaon tagger another event property is included: the number of preselected tagging candidates.

- For the **same-side taggers** the transverse momentum p_T of the signal B, the number of Pile-Up vertices and preselected tagging candidates. In addition, for the SS π, the particle p_T, the impact parameter with respect to the primary vertex, ΔR and ΔQ are also included. In case of the SS kaon, the particle p_T, $\Delta\eta$, $\Delta\phi$ and ΔQ are used.

- For the **vertex charge**, the neural net is a little more complex: it uses the three event properties mentioned above, plus the total number of tracks in the secondary vertex, the mean values of the p_T and of the impact parameter with respect to the primary vertex of the tracks in the secondary vertex. Also the mass of the reconstructed opposite B is used,

which is obtained from the total momentum of the particles in the secondary vertex; in addition the neural net uses the absolute value of the weighted vertex charge.

It is important for the correct behaviour of the *MLP* that the input variables are normalized, having values between $[-1, 1]$. For the single particle opposite side taggers the neural net has 5 inputs (6 for the OS kaon) and an intermediate layer of 8 neurons. The vertex charge neural net has 8 inputs and it is constituted by an intermediate layer of 10 neurons. The same side taggers neural nets contain 7 inputs and an intermediate layer of 10 neurons. All the neural nets have one output neuron which allow us to obtain the probability of the tagger to tag a b or a $\bar{b}$.

The probability of the tagger to be correct $p_i = 1 - \omega_i$ is obtained from the neural net output net_i. For each tagger the dependence of the wrong tag fraction ω_i on the neural net output net_i can be parametrized by a polynomial. The true ω is obtained by binning events in the variable net_i by counting right and wrong tags from MC information or on a control channel. This will correct for possible differences between the true ω and the one estimated by the neural net. For example, in case of the vertex charge tagger: $\omega_{vtx} = a_{vtx} + b_{vtx}\ net_{vtx}$ (see Figure 3.17, left). The parameters a_{vtx} and b_{vtx} are fitted using MC events by looking at the true flavour of the B meson. To take into account non-linearities in the neural net output the parametrization can be different for the different taggers (a polynomial up to the 3th degree can be used). The parametrization used for the different taggers is summarized in Table 3.12.

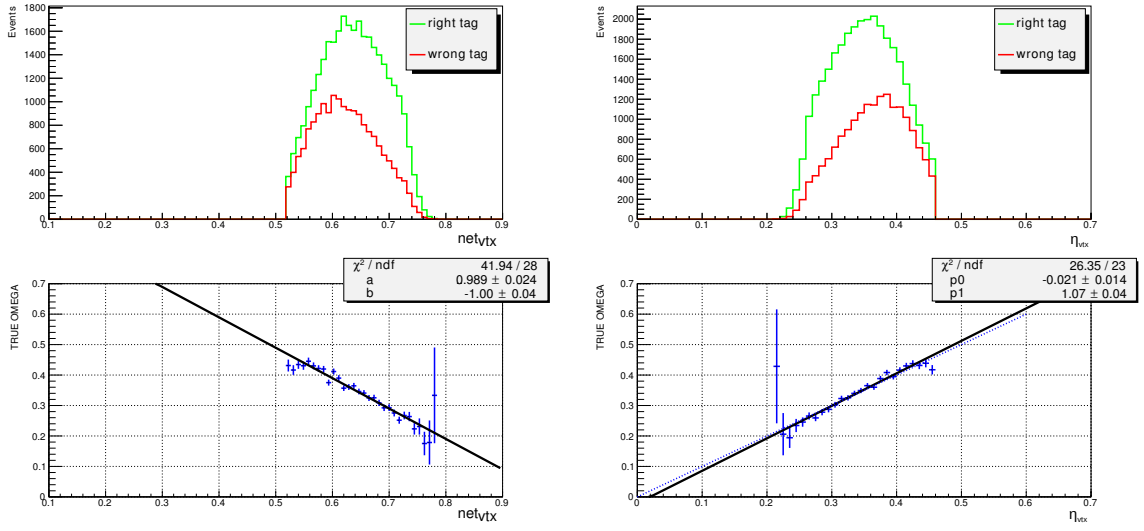


FIGURE 3.17: OS vertex charge tagger for $B^+ \to J/\psi K^+$ MC events. Light (green) is for right-tag, while dark (red) histogram is for wrong-tag assignments. In the second row: Left, the corresponding mistag value ω is fitted from the neural net output (net_i) with a first order polynomial. Right: Measured omega (True Omega) versus the estimated ω using the parametrization on the left.

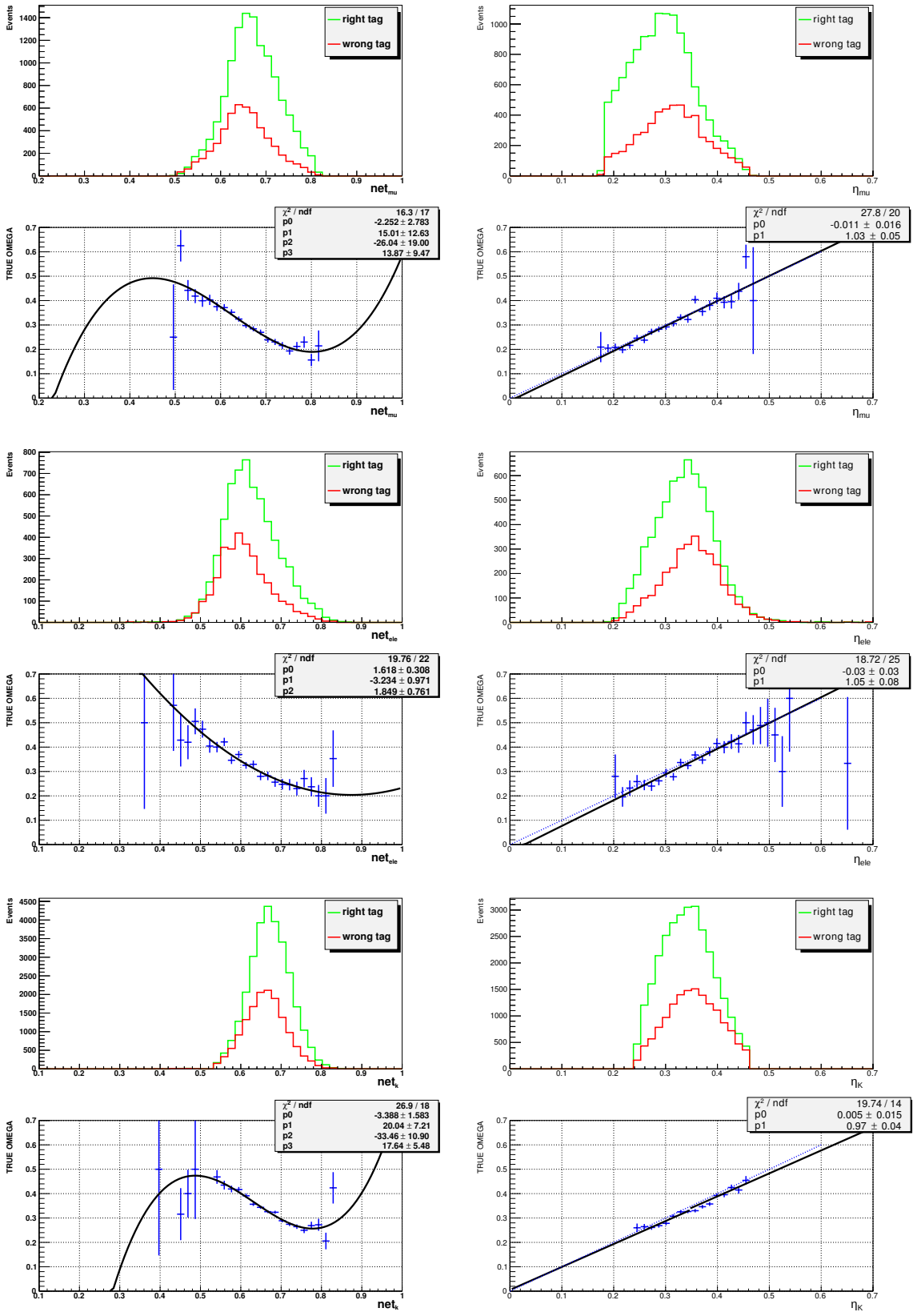


FIGURE 3.18: Single particle taggers (muon, electron and kaon) distributions. Light (green) is for right-tag while dark (red) histogram is for wrong-tag assignments. In the second row: Left, the corresponding mistag value ω is fitted from the neural net output (net_i) with up to 3rd order polynomial. Right: Measured omega (True Omega) versus the estimated ω using the parametrization on the left.

Tagger	a_i	b_i	c_i	d_i
OSμ	-2.25 ± 2.78	15.01 ± 12.63	-26.04 ± 19.00	13.87 ± 9.47
OSe	1.62 ± 0.31	-3.32 ± 0.97	1.85 ± 0.76	–
OSK	-3.39 ± 1.58	20.04 ± 7.21	-33.46 ± 10.90	17.64 ± 5.48
OSvtx	0.99 ± 0.02	-1.00 ± 0.04	–	–
SSπ	1.34 ± 0.17	-2.15 ± 0.55	0.95 ± 0.42	–
SSK	1.22 ± 0.27	-1.63 ± 0.76	0.40 ± 0.51	–

TABLE 3.12: Parametrization of the neural net output ($\omega_i = a_i + b_i net_i + c_i net_i^2 + d_i net_i^3$) to obtain the estimated mistag fraction.

This parametrization transforms the net_i into the wrong tag fraction estimation, hereafter will be defined as η, which is later compared with the measured ω. In Figure 3.17 (left) we show the comparison of the true ω_{vtx} versus net_{vtx} and its parametrization. Figure 3.17 (right) shows ω_{vtx} versus η_{vtx} after the parametrization, any deviation from the straight line would indicate an over or underestimation of ω (see Table 3.13). In Figures 3.18 and 3.19 you can see the neural net parametrization and the measured ω versus the estimated ω for the rest of the taggers.

Tagger	p_0	p_1
OSμ	-0.011 ± 0.016	1.03 ± 0.05
OSe	-0.03 ± 0.03	1.05 ± 0.08
OSK	0.005 ± 0.005	0.97 ± 0.04
OSvtx	-0.02 ± 0.01	1.07 ± 0.04
SSπ	-0.0005 ± 0.013	1.00 ± 0.03
SSK	0.016 ± 0.016	0.95 ± 0.06

TABLE 3.13: Parametrization of the measured omega (true omega) with respect to the estimated omega ($\omega_{measured} = p_0 + p_1 \omega_{estimated}$). Any deviation from the straight line ($p_0 = 0$ and $p_1 = 1$) indicates an over or underestimation with respect to the true omega.

3.7.3 Calibration of the taggers

If the parametrization is correct, the η_i can be used to estimate a wrong tag fraction ω_i per each tagger event per event. As the MC is not expected to represent perfectly the data, it is necessary to correct η_i using data with a control channel. For that, a calibration function is implemented for each tagger i:

$$\omega_i(\eta_i) = p_0^i + p_1^i(\eta_i - \bar{\eta}_i) \tag{3.14}$$

where $\bar{\eta}_i$ is the mean value of the distribution of η_i. This correction allows to introduce a linear correction between data and MC. For MC events it will be $p_0^i = 0$, $p_1^i = 1$ and $\bar{\eta}_i = 0$ as the probability of mistag is already precalibrated. The extraction of p_0^i and p_1^i from data will be explained in Chapter 4. Once the calibrated mistag is obtained, it is possible to use the mistag fraction ω in the CP fits or to combine the different tagger decisions.

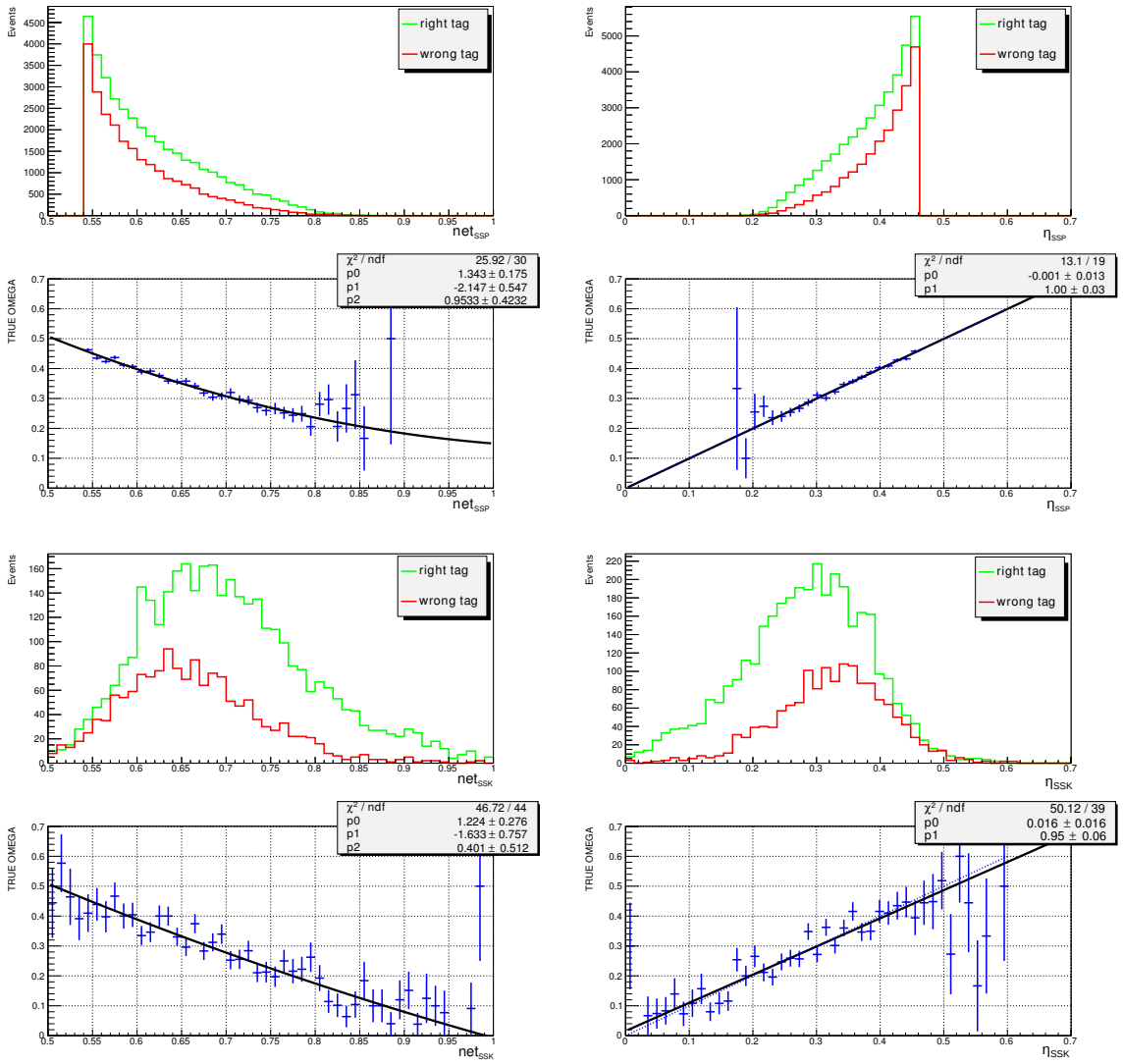


FIGURE 3.19: SS taggers distributions (pion top, kaon bottom). Light (green) is for right-tag while dark (red) histogram is for wrong-tag assignments. In the second row: Left, the corresponding mistag value ω is fitted from the neural net output (net_i) with up to 3rd order polynomial. Right: Measured omega (True Omega) versus the estimated ω using the parametrization on the left.

3.8 Other opposite side taggers

As mentioned before, a number of possibilities have been explored in order to develop new taggers, in particular, for the lepton taggers. Additionally, the reconstruction of the D^0 from the B_{opp}, which can be used complementary to the vertex charge tagger, was developed, also trying to reconstruct the kaon from the opposite side fragmentation chain. The structure of these taggers can be found inside the *FTC* package still in a developmental stage.

3.8.1 Ideas for $b \to c \to l$ tagger

Most of the leptons selected by the muon or the electron tagger come from the $b \to l$ chain. However, some of them could come from the $b \to c \to l$ chain. A schematic view of the process is shown in Figure 3.20.

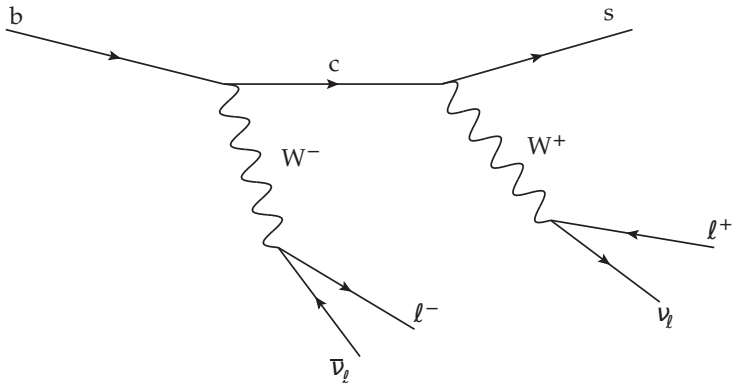


FIGURE 3.20: Semileptonic desintegration chain of the opposite b quark.

For muons and electrons the tagging candidate comes from $b \to l$ decays in 56.7% and 53.2% of the time respectively. However, 14.7% and 13.4% of the muons and electrons come from the $b \to c \to l$ chain giving exactly the wrong answer. Other sources of wrong tag are the misidentified hadrons and other muons in the underlying event. To improve on the wrong tag fraction, a new set of cuts can be introduced to select muons and electrons from the $b \to c \to l$ chain and assign them the corresponding decision. Assuming the same tagging efficiency for $b \to c \to l$ leptons, a reduction of $\sim 10\% - 15\%$ on the wrong tagged events could be easily assumed. This future improvement could roughly increase the tagging performances of the lepton taggers by $20\% - 30\%$.

3.8.2 Opposite side D^0 Tagger

The D^0 opposite side tagger looks for a vertex in the opposite side which can be associated to a D^0 meson decay from the opposite b hadron. If a D^0 is found the B_{opp} would contain a b quark, while if a $\bar{D}^0$ is found the B_{opp} contains a $\bar{b}$ quark. As all the other opposite side taggers, the opposite B can oscillate, so a dilution is unavoidable.

To reconstruct the D^0 meson decaying into $K^+\pi^-$, all the tracks used by the vertex charge tagger are used (the reconstruction of the secondary vertex has been explained in Section 3.3.4). All the possible combinations of 2-tracks are studied, requiring the two tracks to have opposite charge. One of the tracks must be identified as a pion ($\Delta LL_{k-\pi} < 3$ and $\Delta LL_{p-\pi} < 10$) while the other as a kaon ($\Delta LL_{k-\pi} > 6.5$ and $\Delta LL_{k-\pi} - \Delta LL_{p-\pi} > -3.5$). Some extra cuts to maximize the

probability to find a D^0 can be used: only long tracks are used, a cut on the impact parameter significance with respect to the primary vertex of the tracks has to be $IP/\sigma_{IP} > 2.5$ and the p_T of the tracks must be $p_T > 0.3$ GeV/c. The charge of the kaon will directly give the tagging decision. If a K^+ is found a D^0 meson would be tagged while if a K^- is found a $\bar{D}^0$ will be tagged.

In a first step, only 2-track combinations with their reconstructed mass inside a mass window of 400 MeV/c with respect to the D^0 mass was used, but we realized that this cut was not useful at all. If no cut is set, different kind of D meson decays can be reconstructed, increasing the tagging efficiency without increasing the mistag fraction. In Table 3.14 you can see the preliminary performances of the D tagger on MC10 $B^+ \to J/\psi K^+$ events.

D **tagger Performance**			
	ϵ_{tag}	ω	ϵ_{eff}
$D^0 \pm 400$ MeV/c	0.54 ± 0.02	29.68 ± 1.41	0.09 ± 0.02
no mass window	1.86 ± 0.02	29.83 ± 0.60	0.31 ± 0.01

TABLE 3.14: Tagging performance for D taggers (with/without D^0 mass window) on $\sim 500\,000$ $B^+ \to J/\psi K^+$ MC10 events.

Due to the low ϵ_{eff} and the high correlation with the vertex charge tagger and the OSK the tagger was eventually not used. Further studies are needed to understand and possibly reduce the big correlation between the other taggers. Although some strategies can be devised to merge the vertex charge tagger or the OSK with the D tagger. For example, if both are present we can use always the tagging decision with the lower mistag (i.e. the one from the D tagger).

3.8.3 Opposite side Fragmentation kaon tagger

The opposite side fragmentation kaon tagger looks for a kaon coming from the fragmentation chain of the opposite B hadron. This tagger can be used only when a B_s^0 is produced as opposite B, which occurs around $\sim 10\%$ of the times. As for the same side kaon tagger, the opposite side fragmentation kaon is produced with the production of the B_{opp} meson and no dilution is introduced due to B_s^0 oscillations. The charge of the kaon directly indicates the flavour of the B_{opp} at the production time.

The presence of a secondary vertex is required, and the fragmentation kaon is required to be inside a cone of 5 mrad given by the vertex direction with respect the primary vertex. To maximize the probability of a tagging candidate to be a fragmentation kaon from the opposite side, the kaons are selected requiring a momentum $p>$ 3 GeV/c, $p_T>$ 0.4 GeV/c, and an impact parameter with respect to the primary vertex with and without significance $IP < 1.5$ mm and $IP/\sigma_{IP} < 2$. In addition, the kaon has to be identified with $\Delta LL_{k-\pi} > 0$ and

$(\Delta LL_{k-\pi} - \Delta LL_{p-\pi}) > -1$. If more than one candidate are selected the one with highest p_T is used.

The tagging performance ϵ_{eff} of the fragmentation kaon is around 0.1% and has a high ω ($\sim 40\%$). The reason for this is that the fragmentation kaon can be found in only B_s^0 mesons (10%) which produce a charged kaon (50%), but only if a secondary vertex is reconstructed (16%). One also has to take into account the kaon particle identification efficiency, which leads to a effective efficiency of $\sim 0.15\%$. Further developments could be introduced to increase the performance, in particular in the selection of the B_{opp}, or simply omitting the cone cut around the reconstructed secondary vertex.

3.9 Conclusions

In this chapter all the flavour tagging algorithms of the LHCb, opposite side and same side, have been explained in detail. To sum up, we have described how the tagging decision and the mistag estimation for each individual tagger are obtained. The different variables used in each of the taggers have been detailed as well. However, it is important to mention that the optimal values that maximize the performances of the taggers in this chapter had been obtained in MC10. So, a further optimization and calibration would be needed when running on real data, as detailed in Chapter 4.

Chapter 4

Optimization and Calibration

Flavour Tagging algorithms were initially developed on MC events. However, there are differences between data and MC that could make the tagging response not reliable. That is why the optimization procedure is applied on LHCb data, with the aim to find the set of cuts which maximize the performances. The calibration aims to obtain a reliable mistag estimation which can be used in the CP fits later on. In this chapter, both procedures will be explained in detail.

4.1 Introduction

The flavour tagging algorithms were developed using Monte Carlo (MC) simulation as detailed in Chapter 3. Differences in data and MC or even between data in different running conditions, can make the tagging algorithms response unreliable in data. Data-based optimization and calibration processes are needed to maximize the tagging performances and to obtain a reliable per-event mistag estimation. To do that, several flavour-specific control channels were studied. In the case of the OS and SSπ taggers the $B^+ \to J/\psi K^+$, $B^0 \to D^{*-}\mu^+\nu$, $B^0 \to J/\psi K^*$ and $B^0 \to D^-\pi^+$ channels were used.

The SS kaon can only be used in B_s mesons where an extra s quark is available in the fragmentation chain to form a kaon. From MC studies, the introduction of the same side kaon (SSK) tagger is expected to almost double the global tagging power [87]. It is therefore of special interest in all analyses that need to tag a B_s meson, like for example the ϕ_s measurement with the $B_s^0 \to J/\psi\phi$ channel. The optimization and calibration of the SSK was performed using $B_s^0 \to D_s^-\pi^+$ data.

All these channels were used for the optimization, calibration and measurement of the tagging performances of the opposite-side and same-side taggers for different sets of data collected with the LHCb detector before the 2012 physics run. The first 35 pb^{-1} were collected during the

2010 physics run and allowed us to perform the first optimization and calibration on data. The results will be labeled as *Moriond 2011* in Tables and text. In summer 2011, 370 pb^{-1} became available and a second optimization and calibration were performed. The corresponding results will be labeled as *LeptonPhoton 2011*. Finally, 1 fb^{-1} was collected by the end of 2011 leading to the results labeled as *Moriond 2012*.

4.2 Control channels

Control channels are decays to flavour-specific final states. The flavour of the signal B is given by its decay products and so that it is possible to determine the wrong tag fraction directly from data by comparing the known flavour of the B meson with the one given by the tagging algorithms, or by fitting the flavour oscillation in B^0 mesons.

In B_u channels, like the $B^+ \to J/\psi K^+$, the SSπ from the fragmentation chain can be, in some of the cases, a true kaon due to misidentification, as a $s\bar{s}$ pair can be created instead of a $d\bar{d}$ pair. This could cause biases in the optimization or in the calibration of the SSπ, because of the cut on the $\Delta LL_{K-\pi}$ distributions. To be sure that only true pions are identified as such, ideally only B_d channels should be used.

Various control channels are used for tagging. A summary of the fits to disentagle signal from background and to fit the flavour asymmetry in case of B^0 mesons, together with the selection criteria, can be found in the Appendix A. In the following, a short description is done for the channels used in the optimization and calibration procedures. The $B^0 \to J/\psi K^*$ channel is described in Appendix A as it is only used to measure the performances as a crosscheck. In Table 4.1 the different yields are shown for the different channels and luminosities used in the different steps of the optimization and calibration process.

	35 pb^{-1}	370 pb^{-1}		1 fb^{-1}	
$B^0 \to D^{*-}\mu^+\nu_\mu$	$\sim$ **48 000**	$\sim$ 482 000	$(B/S \sim 0.14)$	–	
$B^+ \to J/\psi K^+$	$\sim$ 11 000	$\sim$ **85 000**	$(B/S \sim 0.035)$	$\sim$ **251 000**	$(B/S \sim 0.034)$
$B^0 \to J/\psi K^*$	$\sim$ 3 300	$\sim$ 33 000	$(B/S \sim 0.29)$	$\sim$ 107 000	$(B/S \sim 0.40)$
$B^0 \to D^-\pi^+$	–	–		$\sim$ 105 000	$(B/S \sim 0.04)$

TABLE 4.1: Yields available and B/S for different luminosities on different control channels. In bold the channel used in each optimization.

4.2.1 $B^0 \to D^{*-}\mu^+\nu_\mu$ channel

Semileptonic B decay channels are of wide interest in the LHCb due to their high yield. In particular, the $B^0 \to D^{*-}\mu^+\nu_\mu$ channel is the flavour specific mode with the highest yield,

and it can be used for tagging optimization purposes. The reconstruction of events with a missing neutrino is difficult at an hadronic collider with a non hermetic detector. However, given the good performances of LHCb in vertex reconstruction and particle identification, these type of measurements are still feasible, and a good selection efficiency can be achieved. For the $B^0 \to D^{*-}\mu^+\nu_\mu$ channel, the background contamination is limited and the background to signal ratio B/S is measured to be ~ 0.14 in the signal mass region [88]. The measurement of the decay time is calculated using the measured B^0 decay length, the reconstructed B^0 momentum and a correction for the missing neutrino, known as "k-factor", determined from simulation and parametrized as a function of the reconstructed B^0 invariant mass. More details on the selection criteria and the "k-factor" can be found in Appendix A.

The flavour of the reconstructed signal B^0 is given by the charge of the μ which exactly determines the flavour of the meson at decay time. If a μ^+ from the signal decay products is found, the reconstructed signal is a B^0, while if a μ^- is reconstructed, it will be a $\bar{B}^0$. The flavour of the B^0 at generation is determined by the tagging decision. The mixing can be determined by comparing the flavour of the reconstructed B^0 and the flavour indicated by the tagging decision. To obtain the mistag fraction ω and the tagging performances a fit to the flavour oscillation as a function of the proper-time is necessary. This fit is also detailed in Appendix A.

4.2.2 $B^+ \to J/\psi K^+$ channel

The $B^+ \to J/\psi K^+$ is a self-tagging channel which can be used to measure the tagging performance of the different taggers and to perform the optimization and calibration of the tagging algorithms. The charge of the kaon coming from the decay products of the B will unambigously indicate its b quark content. If a K^+ is reconstructed, a $\bar{b}$ quark ($d = +1$) will be tagged at decay time, while if a K^- is found a b quark ($d = -1$) will be tagged. As charged mesons do not oscillate, the flavour at production is exactly the same as the flavour at the decay time. Thus, the quark component of the reconstructed signal B can be directly compared with the estimated quark component of the B meson at the production point, given by the flavour tagging algorithms, obtaining the tagging performances.

The $B^+ \to J/\psi K^+$ channel is used for opposite-side tagging studies. Its high yield, and its similarity to the $B_s^0 \to J/\psi\phi$ channel, which can be used to measure the ϕ_s phase, makes it one of the best control channels in LHCb. The event topology, the trigger and the data selection are in fact very similar between these two channels and consequently the tagging performances of the opposite side taggers can be safely assumed identical [89].

The $B^+ \to J/\psi K^+$ candidates are selected by combining $J/\psi \to \mu^+\mu^-$ and K^+ mesons candidates. To enhance the sample of signal events and reduce the dominant background contamination from prompt J/ψ mesons with random kaons, only the events with a reconstructed decay

time of the B candidate $t > 0.3$ ps are selected. The signal events are statistically disentangled from the background, which is dominated by partially reconstructed b hadron decays ($J/\psi X$), by means of an unbinned maximum likelihood fit to the reconstructed B^+ mass and decay time. With 370 pb^{-1} of data, a total of $\sim$ 85 000 signal events were selected with a background to signal ratio $B/S \sim 0.035$ calculated in a window of ± 40 MeV/c^2 centered around the B^+ mass. The selection of this channel and the full fit are detailed in Appendix A.

4.2.3 $B^0 \to D^-\pi^+$ channel

The $B^0 \to D^-\pi^+$ channel is a flavour specific channel as the charge of the π (or of the D meson) determines the flavour of the B meson at the decay time which can be compared with the flavour at generation given by the tagging algorithms. In order to measure the mistag fraction and the tagging performances a fit to the flavour oscillation as a function of the proper-time can be made.

The selection of the $B^0 \to D^-\pi^+$candidates is described in [90]. In particular, only events with a reconstructed decay time $t > 0.2$ ps and triggered by one of the different topological trigger lines were used.

This channel is used since the end of 2011 for optimization and calibration purposes, when with 1 fb^{-1} of data around 105,000 $B^0 \to D^-\pi^+$ signal events, with a background to signal ratio $B/S \approx 0.04$, were collected.

4.2.4 $B_s^0 \to D_s^-\pi^+$ channel

For the SSK, B_s channels must be used for the optimitzation and calibration. The $B_s^0 \to D_s^-\pi^+$ channel is a flavour-specific mode, and the tagging performances can be obtained by a fit to the flavour oscillation as a function of the decay time. The oscillation is much faster in B_s^0 ($\Delta m_s = 17.725 \pm 0.041 \pm 0.026$ ps^{-1} [91]) than in B^0 channels. The trigger and stripping lines, as well as the selection that was used to select the B_s decays, are the same used in the Δm_s analysis. Together with the full fit to the flavour oscillation, they are described in [92].

4.3 Optimization of Flavour Tagging algorithms

The FT algorithms are based on a set of cuts on certain properties of the tagging candidates which allow us to select the best tagging particle. Given that data are not perfectly described by MC, differences between both samples are expected. Thus, it is not necessarily true that the

set of cuts optimized for the best tagging performance on MC are also optimal on data, which makes a new optimization necessary.

Figure 4.1 shows some of the most important discrepancies in the variables used for tagging in data and MC. Also, differences can appear between several data sets collected with different running conditions or different reconstruction versions, as they can introduce corrections on some of the measured quantities. In Figure 4.2 differences between the track multiplicity, the number of primary vertices and the signal B p_T are shown. These last 3 variables can also play an important role in the calibration process detailed in Section 4.4.

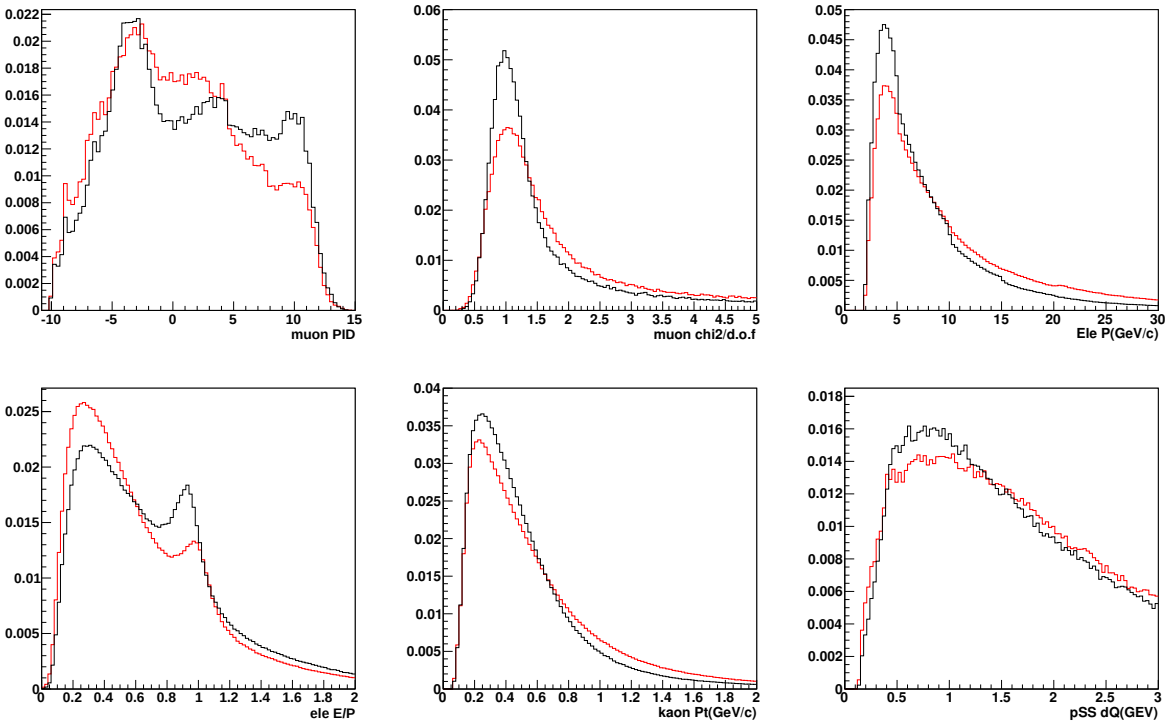


FIGURE 4.1: Normalized distributions of tagging candidates for $B^+ \to J/\psi K^+$ signal events. Red line is data and black line corresponds to simulated MC10 events. From top left to bottom right: $\Delta LL_{\mu-\pi}$, μ $\chi^2/d.o.f$, electron p and E/p, kaon OS p_T and π SS ΔQ.

The optimization procedure aims at finding the set of cuts to be used in the selection of the taggers which will maximize the ϵ_{eff}. It consists of a discrete scanning of each cut on a given variable, keeping the other fixed, and finding the value of the variable that maximizes the tagging power. For each set of cuts, the ϵ_{eff} is computed from the tagging efficiency ϵ_{tag} and mistag fraction ω. The ϵ_{tag} is determined by counting the events where the taggers give an answer, and the ω is determined by counting the right and wrong tagged events or by fitting the flavour oscillation in case of neutral B^0 mesons.

In the following sections the different optimization steps which have been used to optimize the LHCb flavour tagging algorithms are explained. At first, the optimization was performed

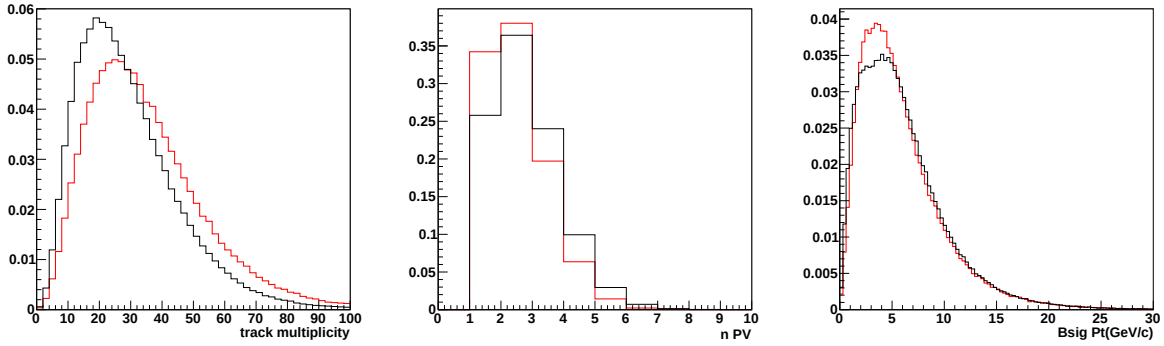


FIGURE 4.2: From left to right: Distributions of track multiplicity, number of primary vertices and p_T of the signal B meson for $B^+ \to J/\psi K^+$ signal events. Red line is data and black line is simulated MC10 events.

using the $B^0 \to D^{*-}\mu^+\nu_\mu$ channel. Once the statistics became sufficient (370 pb^{-1}), $B^+ \to J/\psi K^+$ events were used. Additionally, in Section 4.7.1, an optimization procedure which could be used with high statistics on $B^0 \to D^-\pi^+$ channel is described.

4.4 Calibration of Flavour Tagging algorithms

As discussed in Chapter 3 for each tagger i the estimated mistag probability η_i is calculated from the output of the neural net (e.g. $\eta_i = a_i + b_i net_i$). This information is used for the combination of taggers to calculate the corresponding probability of the B meson to contain a b or a $\bar{b}$ quark. This combination is explained in Chapter 5. In order to have the best possible combination of taggers and have a reliable tagging information, it is necessary that the estimated mistag probability of each tagger and of the combination is correctly calibrated.

In the calibration process, for each individual tagger i we compare the measured mistag fraction ω with the estimated mistag η, given by the flavour tagging algorithms. Then a correction function is extracted by the linear dependence between ω and η:

$$\omega(\eta) = p_0 + p_1 \cdot (\eta - <\eta>) \,, \tag{4.1}$$

where p_0 and p_1 are free parameters and $<\eta>$ is the mean mistag probability. In order to reduce the correlation between p_0 and p_1 the linear dependence on η has been centered around the average value $<\eta>$. Deviations from $p_0 =<\eta>$ and $p_1 = 1$ indicate that the mistag probability is not calibrated. The linearity assumption is motivated by the fact that the calibration should be a minor correction to account for the differences between data and simulation because η is already pre-calibrated on MC as detailed on Section 3.7.2.

The calibration of the individual taggers in principle should not modify the average measured performances of the corresponding tagger. The calibration simply ensures that the value of the mistag probability for a given event is unbiased. However, due to the cuts applied to the probability of the tagging decision to be correct, which is also optimized after the calibration, differences are expected in the tagging performance before and after the calibration procedure in the OSK, vertex charge and SS taggers.

The mistag probability for OS taggers has been calibrated using the control channel $B^+ \to J/\psi K^+$, first on 35 pb^{-1} of data collected in 2010, then with 370 pb^{-1} from the 2011 physics run and finally at the end of 2011, when 1 fb^{-1} of data was available. With 1 fb^{-1} the SSK calibration could be performed with $B_s^0 \to D_s^- \pi^+$ and a new calibration with B^0 events is ongoing for the SSπ.

When the $B^+ \to J/\psi K^+$ channel is used, in order to extract the calibration parameters p_0 and p_1, an unbinned maximum likelihood fit to the mass, the tagging decision and η is performed as described in Appendix A. In this fit the tagging part of the signal is written as:

$$\mathcal{P}^{\text{tag}}(r,\eta) = \begin{cases} \epsilon_{tag}\,(1-\omega(\eta))\,\mathcal{P}(\eta) & \text{if r=1, right tag,} \\ \epsilon_{tag}\,\omega(\eta)\mathcal{P}(\eta) & \text{if r=-1, wrong tag,} \\ 1-\epsilon_{tag} & \text{if r=0, untag.} \end{cases} \tag{4.2}$$

where $\omega(\eta) = p_0 + p_1 \cdot (\eta - <\eta>)$. In particular, the probability of mistag η is an observable and can be used to extract the p_0 and p_1 calibration parameters. The η distribution is extracted from data selecting the events in the signal mass window for signal events or in the side-bands for background. Depending on the tagger, the probability density function of η is reasonably well described by a Gaussian, otherwise it is taken as *RooKeyPdf*. To determine the calibration parameters, each tagger is fitted independently. The results are considered only if the fit converges and MINUIT can calculate a positive definite covariance matrix.

A final tagging decision combining all the OS taggers can be obtained and used for the CP fits. In that case, a calibration on the OS combination is also applied to correct for a bias due to correlation effects. This process is detailed in Chapter 5. In the next sections the optimization and calibration procedures are described as they were performed with 35 pb^{-1}, 370 pb^{-1} and 1 fb^{-1} in the individual taggers.

4.5 Optimization and calibration with 35 pb^{-1}

With 35 pb^{-1} of data from the 2010 physics run the first optimization and calibration of the flavour tagging algorithms took place at LHCb [86]. Due to the lack of statistics of other control

channels, the $B^0 \to D^{*-}\mu^+\nu_\mu$ was used for the optimization of the OS and SSπ. For the SSK there was not enough data of any B_s control channel, however, an alternative channel with a high yield could be used; the $D_s^+ \to \phi\pi^+$ decay.

4.5.1 Optimization on $B^0 \to D^{*-}\mu^+\nu_\mu$ events

In the $B^0 \to D^{*-}\mu^+\nu_\mu$ channel, the tagging parameters ϵ_{tag} and ω can be obtained through a fit to the flavour oscillation of the B^0 mesons as a function of the decay time given by the time dependent asymmetry:

$$A(t) = \frac{N^{\text{unmix}}(t) - N^{\text{mix}}(t)}{N^{\text{unmix}}(t) + N^{\text{mix}}(t)} \tag{4.3}$$

where N^{unmix} (N^{mix}) is the number of tagged events which have not (or have) oscillated at the decay time t. In absence of background, the asymmetry is maximal at t=0, and can be defined as:

$$A(t) = (1 - 2\omega)\cos(\Delta m_d t) \tag{4.4}$$

where Δm_d is the B^0–$\bar{B}^0$ oscillation frequency. In the ideal case the amplitude of the oscillation would be one. In reality, this amplitude which is diluted by wrongly tagged events is also diluted by the proper time resolution. In presence of background, a fit to the B^0 oscillation has to be applied considering different kinds of background sources as described in Appendix A.

A very long time is required for the fit when the different background components are taken into account. For a fast computation of the ϵ_{eff}, an approximate value of ω can be extracted from a fit to the oscillation $A(t)$ as defined in Eq. 4.4 neglecting the low background level. It was demonstrated in previous MC studies that the cuts which maximize ϵ_{eff} do not differ much from the cuts obtained when taking background into account [78].

For the optimization, the background sources can be disentangled from the signal by using only events within a tight mass window around the $D^* - D^0$ (143.7 MeV/c^2 $< m(K\pi\pi) - m(K\pi) <$ 147.8 MeV/c^2), the reconstructed D^0 ($|m(K\pi) - M_{D^0}| <$ 24 MeV/c^2) and the D^* mass ($|m(K\pi\pi) - M_{D^*}| <$ 24 MeV/c^2), as well as using only events with a proper time $t > 0.2$ ps to exclude prompt events which would increase the contribution of background events leading to an asymmetry $A \simeq 0$ for small decay times. In Figure 4.3 the D^0 and $D^* - D^0$ invariant mass and the decay time distributions for the $B^0 \to D^{*-}\mu^+\nu_\mu$ channel are shown with the different background components.

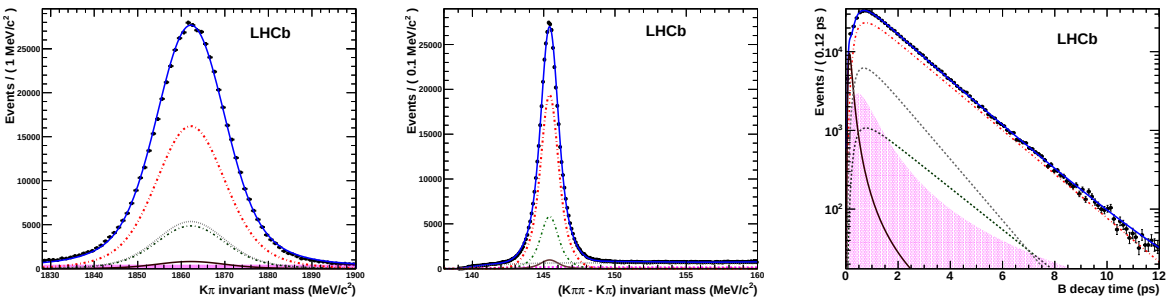


FIGURE 4.3: Distributions of invariant mass $K\pi$ (left), mass difference $m(K\pi\pi)-m(K\pi)$ (center) and decay time t (right) of the $B^0 \to D^{*-}\mu^+\nu_\mu$ events. Black points with errors are data, whereas the blue curve is the fit result. The other lines represent signal (red dot-dashed), $\bar{D}^0$ from B decay background (gray dashed), B^+ background (green short dashed:), D^* prompt background (magenta solid). Magenta filled area: combinatorial background.

However, prior to any asymmetry fit, a simplified approach was used for a pre-optimization. As the $B^0 \to D^{*-}\mu^+\nu_\mu$ events have an oscillation frequency $\Delta m_d \simeq 0.5$ ps^{-1}, as a first approximation, events with $0.2 < t < 2.2$ ps can be considered as not oscillated and events with $4.1 < t < 8.0$ ps can be considered as oscillated. With this simplification the effective efficiency can be easily calculated studying the right and wrong tagged events without the need of a more complicated fit to the flavour oscillation. Although a dilution on the measurement of ω (and ϵ_{eff}) will appear due to the presence of background, the cut on a given variable which maximize the performances obtained with this method should be very close to the optimal.

Once we find the set of cuts which maximizes the performances on the above limited decay time window, a fine tuning can be performed by looking at the whole sample, from $0.2 < t < 8.0$ ps. A cosine term is fitted to all the events neglecting bakground in the tuning sample as in Eq. 4.4, with Δm_d fixed to the PDG value [6], obtaining an ω which can be used to compute, together with the ϵ_{tag}, the ϵ_{eff} of the sample. In Figure 4.4 the asymmetry distribution as a function of proper-time and the fit to the $A(t)$ as defined in Eq. 4.4 is shown for the individual taggers.

To avoid over-tuning, the data set was randomly split into two samples, one used for tuning, where the optimization was done, and a second used for test. The test sample is used to check if there is a real improvement in the performance with the new set of cuts. The optimization is later on validated using the fit that determines the tagging performances in presence of background.

In Figure 4.5 the full fit to the time-dependent asymmetry can be seen for the first 35 pb^{-1} of data. The superimposed red curve represents the fit function as defined in Appendix A, obtained fixing $\Delta m_d = 0.507$ ps^{-1}. Due to the presence of large (symmetric) background the measured asymmetry reduces at $t \sim 0$. This has been the first evidence of flavour oscillations at LHCb.

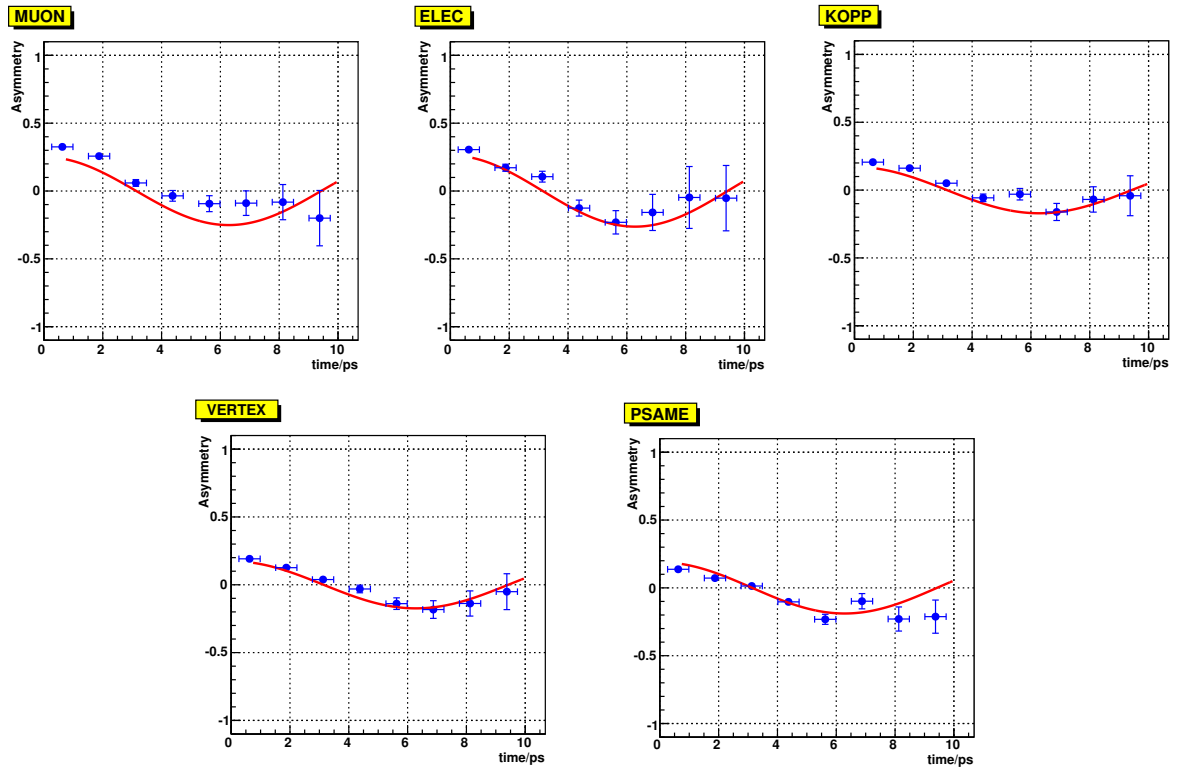


FIGURE 4.4: Asymmetry distributions as a function of proper-time and fit to the function $(1-2\omega)\cos(\Delta m_d t)$ of selected and triggered $B^0 \to D^{*-}\mu^+\nu_\mu$ events in the selected signal region neglecting background and fixing $\Delta m_d = 0.507 ps^{-1}$. From top left to bottom right: Muon tagger, Electron tagger, OS kaon tagger, OS vertex tagger and SS pion tagger.

In spite of the limited statistics, a total sample consisting of 48 k signal $B^0 \to D^{*-}\mu^+\nu_\mu$ events were sufficient to optimize the tagging algorithms. The set of cuts is shown in Tables 4.17, 4.18 and 4.19 labeled as *Moriond 2011*. An additional set of cuts for the opposite-side muon and kaon taggers was obtained with a parallel optimization procedure using *sWeights* [93] on the $B^+ \to J/\psi K^+$ channel [86]. Only the set of cuts for the OS kaon tagger performed better than the one obtained with the $B^0 \to D^{*-}\mu^+\nu_\mu$ channel and were chosen by default in the *FT* package.

4.5.2 Calibration on $B^+ \to J/\psi K^+$

With 35 pb^{-1} of data 11 k $B^+ \to J/\psi K^+$ events could be used for the calibration of the OS and SSπ taggers. In Figure 4.6 we show the η distributions for the right and wrong tagged events and the histogram of the ratio $W/(W+R)$ for signal and background events. From the signal region ($|m - M_{B^+}| < 40$ MeV/c^2) the calibration parameters are extracted, and a linear dependence between ω and η is confirmed, while for the background, extracted from the sidebands, it is verified the independence of the measured mistag fraction ω from the calculated

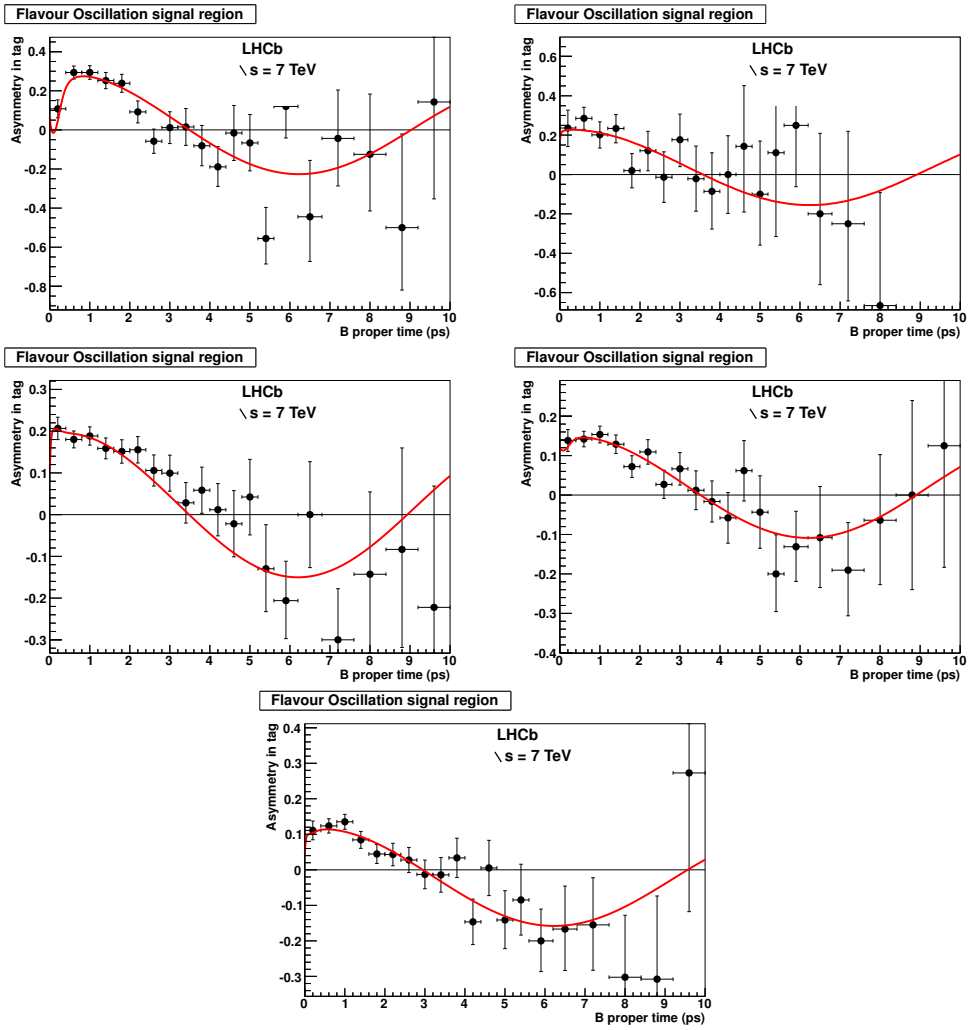


FIGURE 4.5: Time dependent asymmetry measured in the $B^0 \to D^{*-}\mu^+\nu_\mu$ channel for the events in the signal mass region, as determined by using the tagging decision of single taggers and by the combination of all the taggers. Only 35 pb^{-1} of data from the 2010 physics run were used. The red curves correspond to the full fit result obtained by fixing $\Delta m_d = 0.507$ ps^{-1}. Top left: Muon tagger, top right: Electron tagger, middle left: opposite side kaon tagger(*), middle right: opposite-side vertex tagger, bottom: same-side pion [86].

mistag probability η. The picture corresponds to the OS kaon tagger and the cuts from Table 4.17. Similar plots can be obtained for the other taggers.

The calibration was performed with a tuning of the neural nets based on a previous version of the MC10 data, and due to the low statistics, the whole data sample was used. The η distribution was described by a Gaussian with $<\eta>$ left free in the fit (muon, electron, kaon) or a by *RooKeysPdf* with fixed $<\eta>$ (for all the other cases). A summary of the parameters of the fit relevant for the calibration are reported in Table 4.2.

From the Table 4.2, one can notice that the most relevant corrections apply to the vertex charge tagger, where both p_1 and $p_0 - <\eta>$ deviate significatively from 1 and 0. The parameter p_1 is

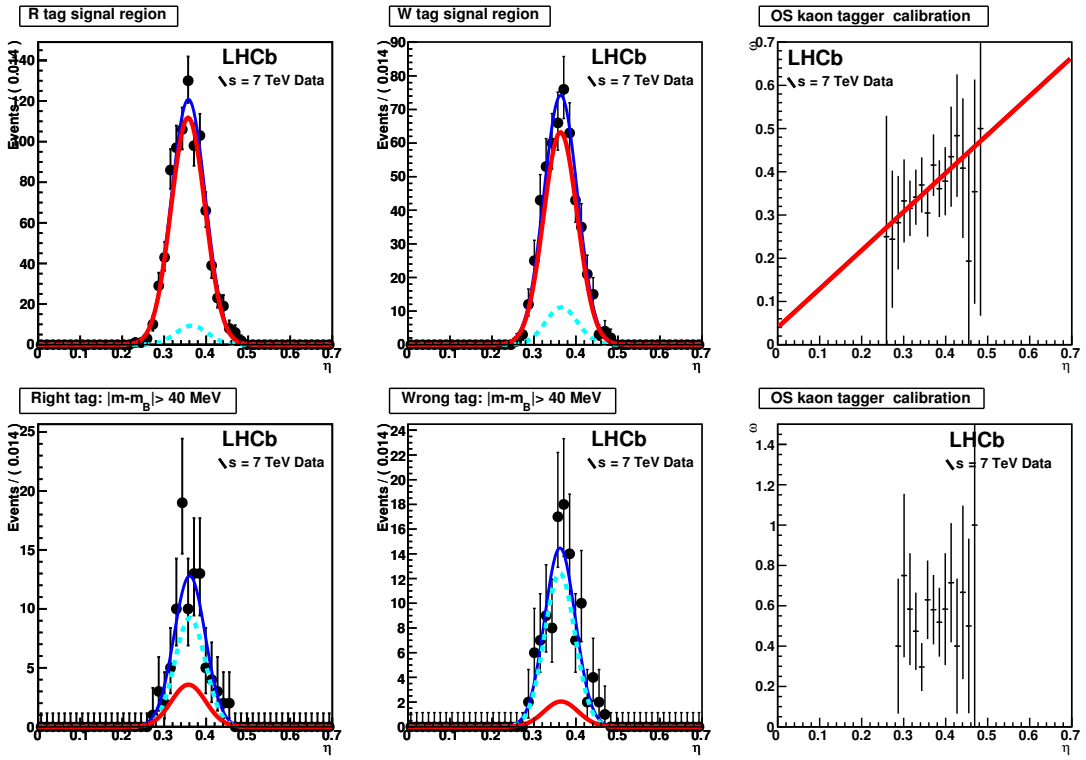


FIGURE 4.6: OS kaon (*) η distributions for the right (left) and wrong (middle) tagged events in the signal (background) mass window $|m - M_{B^+}| < 40$ ($|m - M_{B^+}| > 40$) MeV/c^2 on the top (bottom). Right plot: the histogram of the ratio $W/(W + R)$ for the background subtracted events in the signal mass window. Points with errors are the data, blue line is the fit result, the red and dashed light blue lines are the signal and background components respectively. The linear function represent the result of the mistag probability calibration for the signal.

35 pb^{-1} of $B^+ \to J/\psi K^+$ signal events			
Taggers	p_0	p_1	$< \eta >$
μ	0.311±0.020	1.36±0.31	0.316
e	0.290±0.034	1.47±0.40	0.341
K	0.359±0.014	0.89±0.34	0.359
$Q_{\rm vtx}$	0.451±0.012	0.36±0.16	0.398 (fix)
SSπ	0.425±0.011	0.89±0.21	0.415 (fix)

TABLE 4.2: Calibration parameters for 35 pb^{-1} determined using $B^+ \to J/\psi K^+$ signal events. The quoted errors are statistical.

generally affected by a large error, while the parameter p_0 is better determined. The precisions found are in agreement with the results of the studies performed on simulated data that can be found in the LHCb note [89].

As expected, after the calibration, the parameter p_1 becomes compatible with 1 also in the cases it was not before, while the difference $p_0 - < \eta >$ becomes compatible with 0, confirming that the calibration succeeded. The results on the calibration parameters after the calibration

procedure are summarized in Tables 4.3, where the calculated mistag probability, corrected according to the calibration procedure, is labeled as η_c.

35 pb^{-1} of $B^+ \to J/\psi K^+$ signal events			
Taggers	p_0	p_1	$< \eta_c >$
μ	0.311±0.002	1.00±0.22	0.311
e	0.295±0.003	1.02±0.38	0.293
K (*)	0.361±0.014	1.03±0.39	0.359
Q_{vtx}	0.450±0.013	1.02±0.43	0.446 (fix)
SSπ	0.425±0.012	1.12±0.26	0.421 (fix)

TABLE 4.3: Calibration parameters obtanied with $B^+ \to J/\psi K^+$ signal events, after the correction for the calibration of the individual taggers of Table 4.2. The error is statistical.

4.5.3 Optimization of the SSK tagger

Due to the lack of statistics the optimization of the SSK was performed on $D_s^+ \to \phi\pi^+$ decay as detailed in [94] to exploit the D_s fragmentation. However, the D_s mesons have a much lighter mass than the B_s mesons, which translates to different kinematics. Despite these differences, several MC studies for both $D_s^+ \to \phi\pi^+$ and $B_s^0 \to D_s^-\pi^+$ were done [94] to assess the feasibility of the optimization with this channel. It can be shown that the relevant kinematic properties of the fragmentation kaon are very similar, leading to the possibility of optimizing the tagging cuts using $D_s^+ \to \phi\pi^+$ decays. The tagging variables that were different were $|\Delta\eta|$, $|\Delta\phi|$ and B_{sig} p_T. For this reason, the cuts on these variables were optimized on MC $B_s^0 \to D_s^-\pi^+$.

The selection of $D_s^+ \to \phi\pi^+$ is based on the $B_s^0 \to D_s^-\pi^+$ selection [37]. The optimization procedure is performed for each of the tagging variables, leading to a set of optimized cuts, detailed in Table 4.4, which maximize the tagging performance.

SSK selection cuts (prompt D_s data)
$\Delta LL_{K-\pi} > 4.5$
$\Delta LL_{K-p} >$ -8.5
$p_T > 0.75$ GeV/c
$p > 5.25$ GeV/c
track $\chi^2/d.o.f < 3.75$
$IP/\sigma_{IP} < 4.125$
$\Delta\eta < 0.525$
$\Delta\phi < 0.7$
$\Delta Q < 1.4625$ GeV/c^2

TABLE 4.4: Final set of cuts derived in the optimization procedure on 2010 data with prompt D_s.

After the optimization performed with $D_s^+ \to \phi\pi^+$ data, the performance of the SSK in the $B_s^0 \to D_s^-\pi^+$ decay mode could be preliminarily estimated. Figure 4.7 shows the first oscillation seen in the B_s system, using only the SSK tagger and 580 signal events from the 2010 and 2011 data taking periods. The trigger and stripping lines, as well as the selection that was used to select the B_s decays, is unchanged with respect to the Δm_s analysis [92]. The tagging efficiency for this sample was $16.0 \pm 0.6\%$ and the wrong tag fraction obtained from the asymmetry fit [94], $\omega = 33 \pm 5\%$. Performing the calibration is not appropriate with the $D_s^+ \to \phi\pi^+$ decay due to the mentioned kinematic differences.

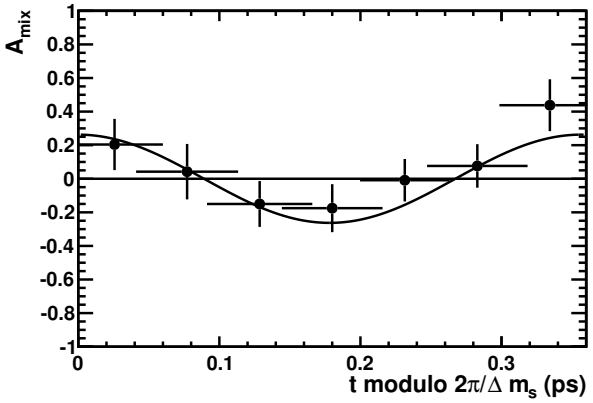


FIGURE 4.7: Mixing asymmetry in $B_s^0 \to D_s^-\pi^+$ as a function of time modulo $2\pi/\Delta m_s$.

4.6 Optimization and calibration with 370 pb^{-1}

Once a sufficient statistics was collected, $B^+ \to J/\psi K^+$ events were used for optimization purposes. Previously, the lack of statistics (only 10 k signal $B^+ \to J/\psi K^+$ events on 35 pb^{-1}) did not allow to perform an optimization of the tagging performances for all the taggers without the risk of over-tuning.

4.6.1 Optimization with $B^+ \to J/\psi K^+$

The tagging efficiency ϵ_{tag} can be determined from the fraction of signal events where the taggers give an answer. As B^+ mesons do not oscillate, the mistag fraction ω is simply measured by comparing the tagging decision obtained from the taggers with the flavour of the signal B at the decay, by counting the number of signal events that are correctly or wrongly tagged.

Since the tagging algorithms perform differently on background events ($\omega \sim 50\%$), the only issue is to carefully disentangle the signal from the background sources. Background can be

easily distinguished from the signal looking at the mass distribution. Figure 4.8 shows the mass distribution of the selected and OS tagged events, together with the superimposed fit.

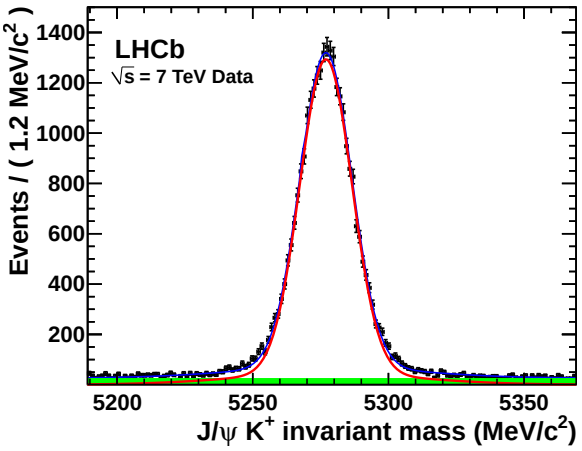


FIGURE 4.8: Mass distribution of OS tagged $B^+ \to J/\psi K^+$ events. Black points are data, the solid blue line, red line and green area are the fit, the signal and the background components, respectively.

The optimization was performed considering all the events in a mass window around the B meson of 25 MeV/c^2 as signal, justified by the low B/S of 0.034. To avoid over-tuning the sample was splitted randomly into a tune sample and a test sample, where the tagging performances were crosschecked. The tagger cuts were iteratively tuned improving the effective tagging power in the training sample. In Figure 4.9 the right and wrong tagged events distribution and its ϵ_{eff} are shown for different variables on different taggers.

The optimization process performed in summer 2011 with $B^+ \to J/\psi K^+$ events confirmed that we were already very near to the optimal set of cuts, obtained with the $B^0 \to D^{*-}\mu^+\nu_\mu$ channel with 2010 data, for most of the taggers. The optimized tagging cuts obtained on 370 pb^{-1} of data from the 2011 physics run are shown in Tables 4.17, 4.18 and 4.19 labeled as *LeptonPhoton 2011*.

Optimization for High Multiplicity and PileUp events

With the data collected in 2010 (35 pb^{-1}), the performances calculated for different numbers of primary vertices were compatible within the statistical error. However, the results seemed to indicate a slight deterioration of the tagging power with increasing number of collisions [86] as indeed it was found in MC (see Chapter 3). With 370 pb^{-1} of data from the 2011 LHCb physics run, the deterioration of the tagging performances with increasing the number of primary vertices was clearly seen. In Table 4.5 the tagging performances for the OS combination and the kaon taggers on $B^0 \to D^{*-}\mu^+\nu_\mu$ events according to the number of primary vertices are shown.

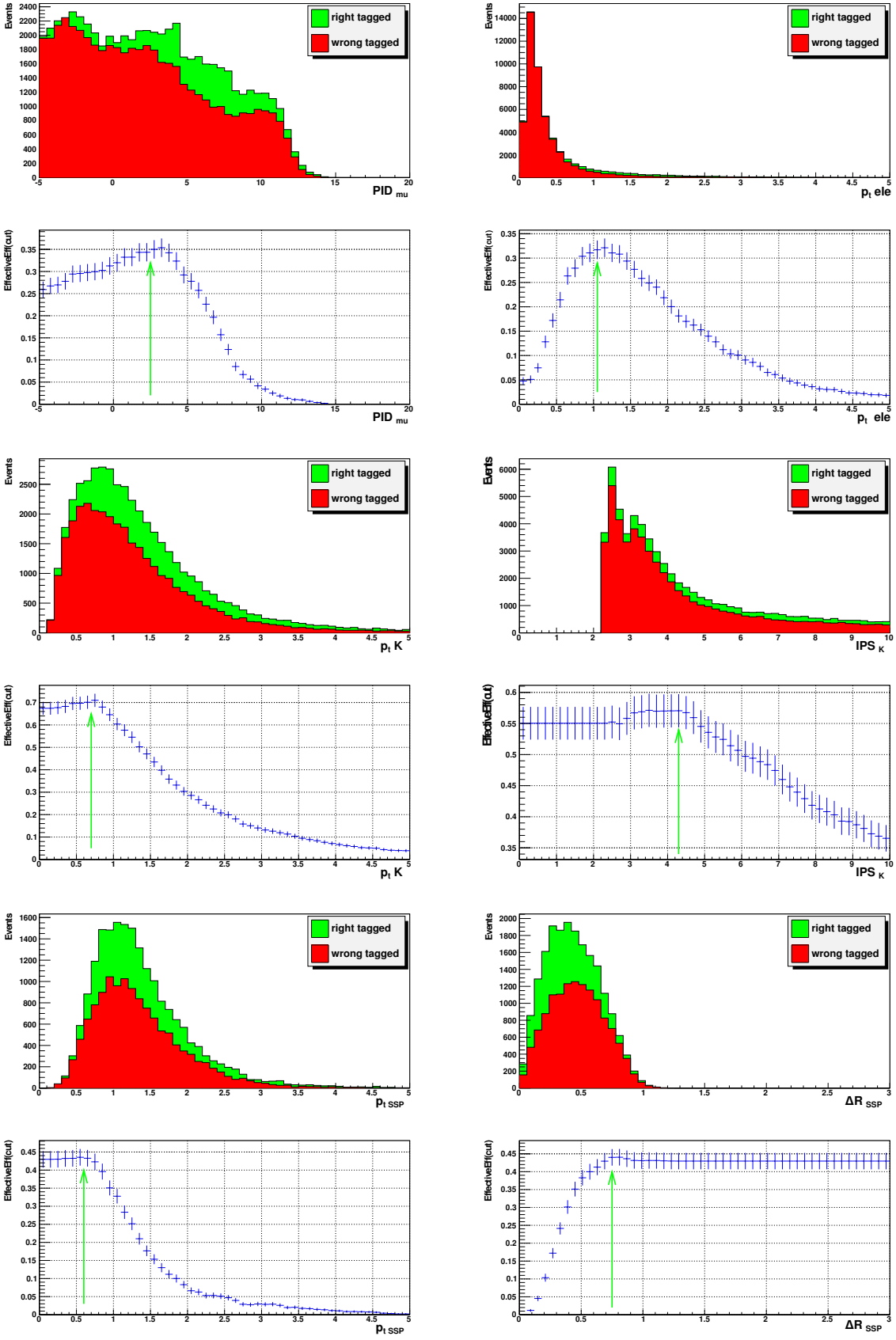


FIGURE 4.9: Above: Right and wrong tagged events for $B^+ \to J/\psi K^+$ events. Below: Calculated effective efficiency as a function of the applied cut. From left to right: muon $\Delta LL_{\mu-\pi}$, electron p_T, OS kaon p_T and IP/σ_{IP} and SSπ p_T and ΔR. The arrow indicates the cut applied.

	PV	$\epsilon_{tag}(\%)$	ω (%)	$\epsilon_{eff}(\%)$
OS	1	22.2 ± 0.1	35.9 ± 0.4	1.76 ± 0.10
	2	20.8 ± 0.1	36.0 ± 0.4	1.64 ± 0.09
	>2	18.4 ± 0.1	36.6 ± 0.5	1.33 ± 0.09
OSK	1	14.04 ± 0.09	37.4 ± 0.5	0.9 ± 0.1
	2	13.29 ± 0.09	38.8 ± 0.5	0.69 ± 0.06
	>2	12.7 ± 0.1	38.8 ± 0.6	0.63 ± 0.07

TABLE 4.5: Dependence of the performance of the combined OS taggers and the OS kaon tagger on the number of primary vertices per event for $B^0 \to D^{*-}\mu^+\nu_\mu$ events. Uncertainties are statistical.

In order to improve the performances for events with a high number of Pile-Up vertices or high track multiplicity, the different variables considered in Section 3.5 were studied by introducing a different set of cuts and calibration parameters for the different situations. Even so, the optimal set of cuts which maximize the ϵ_{eff} turned out to be the same, independently of the number of Pile-Up vertices or track multiplicity.

4.6.2 Calibration with $B^+ \to J/\psi K^+$

The calibration with 370 pb^{-1} was performed as with 35 pb^{-1} but this time the neural nets were based on MC10 as described in Section 3.7.2. The parameters of the calibration are summarized in Table 4.6. After the calibration the parameters p_0 and p_1 were compatible with the linearity assumption. The precisions found were also in agreement with the results of the studies performed on simulated data in the note [89].

370 pb^{-1} of $B^+ \to J/\psi K^+$ signal events			
Taggers	p_0	p_1	$< \eta_c >$
μ	0.307±0.009	1.03±0.14	0.303
e	0.322±0.013	1.19±0.21	0.342
K	0.401±0.005	0.99±0.10	0.369
Q_{vtx}	0.393±0.011	0.85±0.13	0.379

TABLE 4.6: Calibration parameters obtanied with $B^+ \to J/\psi K^+$ signal events for the individual taggers. The error is statistical.

Although it was not suitable to use B_u events for the SSπ calibration, due to the low statistics in B_d channels, the calibration procedure was performed on $B^+ \to J/\psi K^+$ events using a 2^{nd} order polynomial, i.e. $\omega = p_0 + p_1 \cdot (\eta - <\eta>) + p_2 \cdot (\eta - <\eta>)^2$. For a proper calibration of the SSπ, only B_d channels should be ideally used. In Figure 4.10 the measured wrong tag fraction versus the estimated η obtained from the tagging algorithms is shown for the SSπ tagger. In this case, the correction of the η with $p_0 = 0.447 \pm 0.006$, $p_1 = 0.59 \pm 0.07$ and $p_2 = -2.37 \pm 0.44$ fits well the observed data.

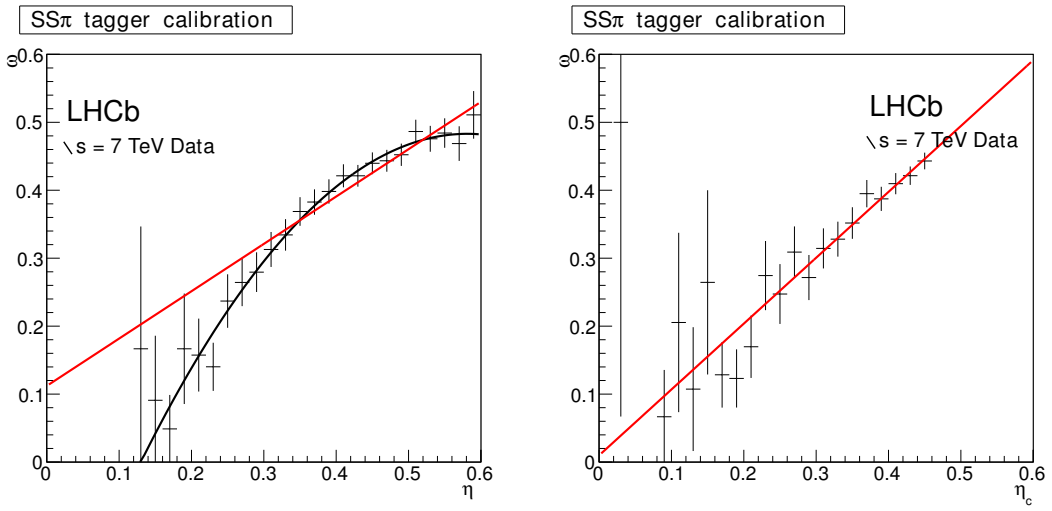


FIGURE 4.10: Measured mistag (ω) versus estimated mistag (η) from the tagging algorithms for SS π tagger before (left) and after (right) the calibration. A 2^{nd} order polynomial is used to calibrate the SSπ.

4.7 Optimization and calibration with 1 fb^{-1}

The optimization and calibration on 1 fb^{-1} of data from the whole 2011 physics run was performed using the $B^+ \to J/\psi K^+$ events for the OS taggers and $B_s^0 \to D_s^- \pi^+$ for the SSK tagger. In case of the SSπ $B^0 \to D^{*-}\mu^+\nu_\mu$ events were used for the optimization while $B^0 \to J/\psi K^*$ and $B^0 \to D^-\pi^+$ events were used for calibration purposes. Additionally an optimization was tried with $B^0 \to D^-\pi^+$ for the OS taggers and the SSπ.

4.7.1 Optimization with $B^0 \to D^-\pi^+$

With enough statistics an optimization with $B^0 \to D^-\pi^+$ signal events can be performed through a fit to the asymmetry. The advantage of the $B^0 \to D^-\pi^+$ channel with respect to the $B^0 \to D^{*-}\mu^+\nu_\mu$ channel is that, in the first one, all the decay products are reconstructed, while the presence of the neutrino makes the reconstruction of the B momentum incomplete in the second channel. In case of the $B^0 \to D^{*-}\mu^+\nu_\mu$ events one has to rely on MC and possible biases can be introduced in the optimization.

This encouraged us to perform the optimization by fitting the B^0 flavour oscillations as in Eq. 4.4 neglecting the background. In Figure 4.11 the mass and decay time distributions for $B^0 \to D^-\pi^+$ events are shown. Again, the statistics available was split into two random samples, one for training and one for testing. To disentangle possible sources of background only events with propertime $t > 0.25$ ps inside a 18 MeV and 25 MeV mass window around the D^- and B^0 invariant mass were used. In Figure 4.12 is shown the flavour oscillation in the selected signal

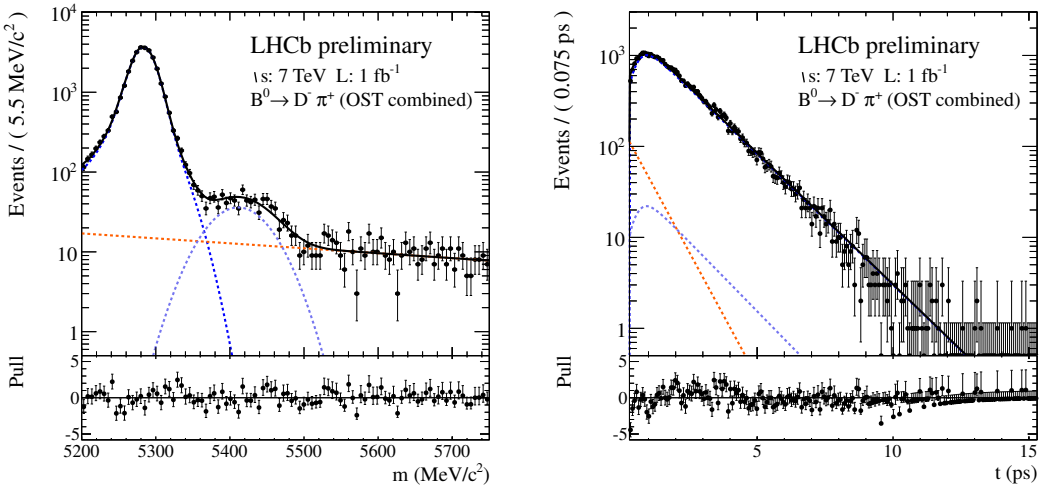


FIGURE 4.11: Mass (left) and decay time (right) distributions for $B^0 \to D^-\pi^+$events. Black points are data, the solid black line is the fit result. The signal component is represented by the blue line, the different background components are: Λ_b (violet) and combinatorial (orange line).

region with the superimposed fit as defined in Eq. 4.4 for all the taggers. A clear oscillation pattern is seen for all the taggers and can be used to find a first set of cuts.

The performances were then calculated with the full fit to the flavour oscillation considering all the background components as defined in A. The performances obtained in the rest of control channels were better with the set of cuts obtained with the $B^+ \to J/\psi K^+$ channel, which were the ones finally used.

4.7.2 Optimization and calibration of the OS taggers

The $B^+ \to J/\psi K^+$ channel was used for optimization and calibration purposes. The performances and the optimal set of cuts were very close to the ones found with 370 pb^{-1} of data. Minor adjustments to the OS electron were applied, as well as to the OS kaon, where a larger tagging efficiency could be achieved. The optimized tagging cuts for the full data sample are defined in Tables 4.10 4.17, 4.18 and 4.19 labeled as *Moriond 2012*. Additionally, with 1 fb^{-1} the tagging performances could be also crosschecked in the $B^0 \to K^+\pi^-$ channel [97].

The calibration of the OS taggers for 1 fb^{-1} was performed with the $B^+ \to J/\psi K^+$ channel. The same MC-based neural net from the 370 pb^{-1} was used. The calibration was first performed splitting randomly the data sample in two and, since values of the fitted parameters were compatible within statistics, this was repeated with the whole data sample. A summary of the parameters of the fit relevant for the calibration are reported in Table 4.7

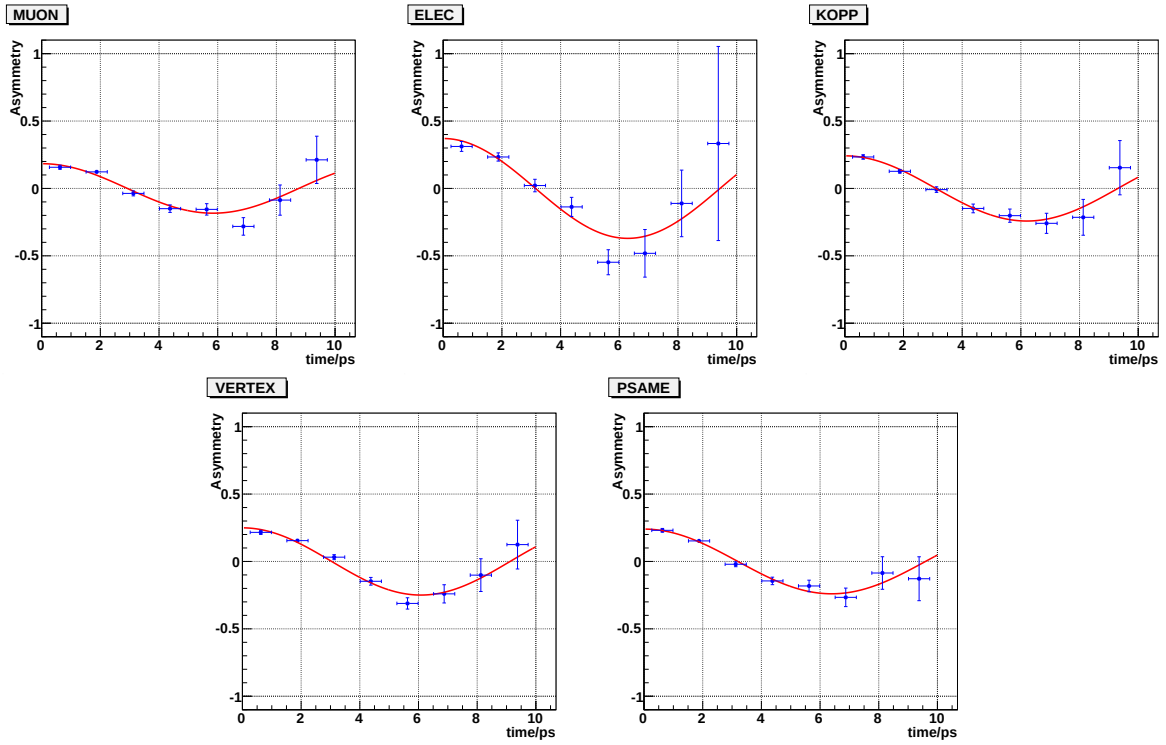


FIGURE 4.12: Time dependent asymmetry measured in the $B^0 \to D^-\pi^+$ channel for the events in the signal mass region and decay time $t > 0.25$ ps. The red curves correspond to the simple fit determined by Eq. 4.4 neglecting backgrounds and fixing the B^0 mass difference to $\Delta m_d = 0.507$ ps^{-1}. From top left to bottom right: Muon tagger, Electron tagger, OS kaon tagger, OS vertex tagger and SS pion tagger.

1 fb^{-1} of $B^+ \to J/\psi K^+$ signal events			
Taggers	p_0	p_1	$< \eta >$
μ	0.309±0.004	1.20±0.06	0.304
e	0.306±0.006	0.974±0.09	0.346
K	0.393±0.002	0.706±0.04	0.354
$Q_{\rm vtx}$	0.404±0.002	0.84±0.03	0.362

TABLE 4.7: Calibration parameters of the OS taggers determined using $B^+ \to J/\psi K^+$ signal events. The quoted errors are statistical.

Additional tests were performed assuming a second order polynomial as calibration curve. In this case, for the OS taggers, the relative error on the quadratic term was above 200%, supporting the linear approximation. After the calibration the values of p_0, p_1 and $< \eta >$ were then remeasured in the whole sample (Table 4.8).

In Figure 4.13, the calibration for the OS kaon performed with 1 fb^{-1} is shown after applying the calibration from Table 4.7. Similar plots can be obtained for the other OS taggers [97]. The calibration parameters were crosschecked for $B^0 \to J/\psi K^*$ and other control channels, confirming that the calibration succeeded [97].

1 fb^{-1} of $B^+ \to J/\psi K^+$ signal events			
Taggers	p_0	p_1	$< \eta_c >$
μ	0.294±0.006	1.044±0.08	0.315
e	0.309±0.009	0.998±0.15	0.307
K	0.394±0.004	0.998±0.10	0.395
$Q_{\rm vtx}$	0.403±0.004	0.992±0.09	0.398

TABLE 4.8: Calibration parameters for the OS taggers obtanied with $B^+ \to J/\psi K^+$ signal events, after the correction for the calibration of the individual taggers of Table 4.7. The error is statistical. The consistency of p_1 to 1 and $p_0 - < \eta_c >$ to 0 indicates the effectiveness of the calibration.

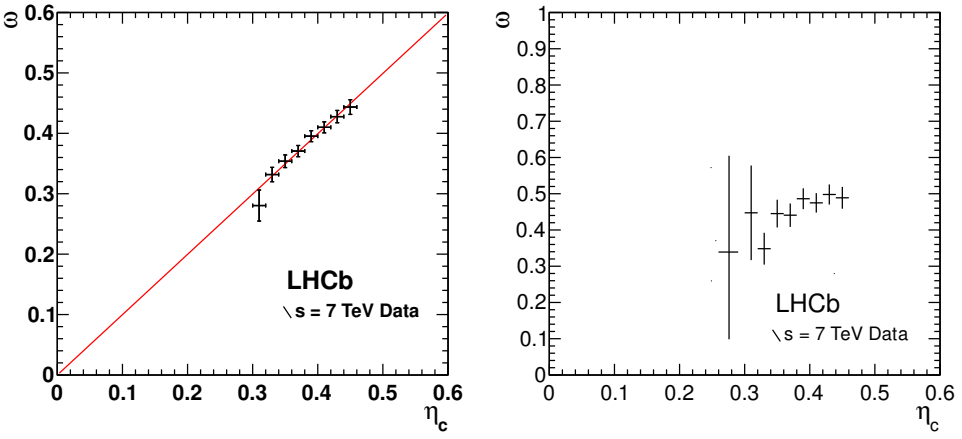


FIGURE 4.13: Calibration plots on OS K tagger for the signal (left) and the background (right) from the fit to the $B^+ \to J/\psi K^+$channel with 1 fb^{-1}. Black points are data, the solid red line is the fit function.

4.7.3 Optimization and calibration of the SSπ tagger

The fragmentation process of a $\bar{b}$ quark to produce B_d^0 or a B^+ are slightly different and can originate a difference both in the performance and in the calibration. The fragmentation of a B_d^0 can produce π^+ π^0 and K^0, while the fragmentation of a B^+ can produce π^-, π^0 and K^-. As the SS tagger in B^+ channels can contain SSπ but also true SSK, we decided to use B^0 channels for the optimization and calibration of the tagger.

The optimization was performed with $B^0 \to D^{*-}\mu^+\nu_\mu$ and two different tunings with similar performances were found. The one that gave the best performance in $B^0 \to D^-\pi^+$ and $B^0 \to K^+\pi^-$ was chosen as reference. In Table 4.18 are reported the cut values (labeled *Morion2012*) corresponding to the optimization.

The calibration of the SSπ tagger is being performed using the $B^0 \to J/\psi K^*$ and $B^0 \to D^-\pi^+$ channels as the one for the single OS taggers. In Table 4.9 are reported the calibration

parameters determined by the fit to $B^0 \to J/\psi K^*$ and $B^0 \to D^-\pi^+$ channels. For comparison, the results from $B^+ \to J/\psi K^+$ channel are shown. Data suggest a better agreement with a second order polynomial. As $B^0 \to D^-\pi^+$ can be used for CP asymmetry measurements, the $B^0 \to J/\psi K^*$will probably be used for the calibration. This work is still ongoing and additional details can be found in [97].

1 fb^{-1} of SSπ				
	p_0	p_1	p_2	$< \eta >$
$B^0 \to J/\psi K^*$	0.434±0.007	0.616±0.086		0.377
	0.443±0.008	0.383±0.129	-2.7±1.1	0.377
$B^0 \to D^-\pi^+$	0.411±0.006	0.573±0.062		0.365
	0.421±0.006	0.337±0.098	-2.84±0.86	0.365
$B^+ \to J/\psi K^+$	0.398±0.003	1.06±0.06		0.378
	0.405±0.005	0.88±0.11	-3.3±1.4	0.378

TABLE 4.9: Calibration parameters for the SSπ tagger determined using different channels. The errors are statistical.

4.7.4 Optimization and calibration of the SSK tagger

With 1 fb^{-1} luminosity the $B_s^0 \to D_s^-\pi^+$ channel could be used to optimize and calibrate the SSK. For the optimization, each cut is in turn varied while all the other cuts are kept at their default values. The tagging performance with the new cut is extracted from the fit of the B_s^0 oscillation. Once all cuts have been scanned this way, the set of cuts which maximize the effective efficiency ϵ_{eff} are chosen. The optimized cuts are summarized in Table 4.10. The performances of the SS kaon tagger after the optimization were evaluated on the test sample giving a tagging efficiency of 16.3±0.4% and a wrong tag fraction, obtained from the asymmetry fit, $\omega = 35.3 \pm 2.1\%$ giving an ϵ_{eff}= $1.4 \pm 0.4\%$. This final set of cuts introduces an absolute enhancement of 0.3% in ϵ_{eff} with respect the cuts found with $D_s^+ \to \phi\pi^+$ data.

For the calibration a required condition is that the estimation of ω as a function of the response of the neural net has to be the same for any B_s decay mode, so that a calibration on $B_s^0 \to D_s^-\pi^+$ events can be used, for instance, for $B_s^0 \to J/\psi\phi$ events. This has been confirmed in MC. The calibration parameters p_0 and p_1 are directly extracted from the unbinned maximum likelihood fit to the B_s^0 decay time distribution where the mistag fraction ω is substituted for the calibration function from Eq. 4.1. Alternatively, the average mistag fraction ω can be determined in bins of the predicted mistag η sorting the events into categories. The corresponding numbers can be found in Table 4.11 and are plotted in Figure 4.14.

As the unbinned maximum likelihood fit is more precise, it is the one used by default, the difference between the two fitters is considered as a systematic error. Other sources of systematics are taken into account and are detailed in Section 6.4.

SSK selection cuts ($B_s^0 \to D_s^- \pi^+$data)
$\Delta LL_{K-\pi} > 3.5$
$\Delta LL_{K-p} >$ -8.5
$p_T > 0.85$ GeV/c
$p > 5.25$ GeV/c
track $\chi^2/d.o.f < 3.75$
$IP/\sigma_{IP} < 4.125$
$\Delta\eta < 0.6$
$\Delta\phi < 0.825$
$\Delta Q < 1.85$ GeV/c^2

TABLE 4.10: Final set of cuts derived in the optimization procedure on 2011 data with $B_s^0 \to D_s^- \pi^+$ events.

	p_0	p_1	$\langle\eta\rangle$
unbinned fit	0.350±0.015	0.51 ± 0.16	0.3237 (fixed)
fit of categories	0.346± 0.015	0.53±0.16	0.3237 (fixed)

TABLE 4.11: Results of the fit of the calibration parameters for an direct unbinned fit of p_0 and p_1 and of a fit in bins of the predicted mistag η.

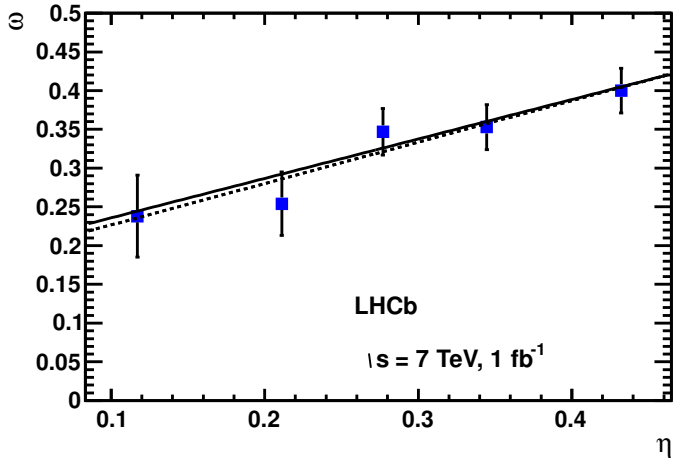


FIGURE 4.14: Average mistag fraction ω in bins of predicted mistag η. The solid line is the result of the unbinned fit for the calibration parameters p_0 and p_1. The dashed line is the result of a linear fit to the data points.

4.8 Conclusions

Due to differences between MC and data, as well as between data from different runs, the set of cuts optimized for one sample can be sub-optimal for another sample we want to analyse. As an example, in Table 4.12 the OS tagging performances are shown for $B^+ \to J/\psi K^+$ events with the optimized cuts on data and the ones obtained from MC.

	cuts	$\epsilon_{tag}(\%)$	ω (%)	$\epsilon_{tag}(1-2\omega)^2$ (%)
μ	*Moriond 2012*	5.20±0.04	30.8±0.4	0.77±0.04
	MC cuts	5.67±0.04	32.5±0.4	0.69±0.04
e	*Moriond 2012*	2.46±0.03	30.9±0.6	0.36±0.07
	MC cuts	2.23±0.03	32.2±0.6	0.28±0.02
K	*Moriond 2012*	17.67±0.08	39.33±0.24	0.81±0.04
	MC cuts	13.09±0.07	39.17±0.24	0.61±0.03
$Q_{\rm vtx}$	*Moriond 2012*	18.46±0.08	40.31±0.24	0.70±0.04
	MC cuts	23.31±0.08	41.94±0.18	0.61±0.03

TABLE 4.12: OS tagging performance for 1 fb^{-1} on $B^+ \to J/\psi K^+$ channels with the cuts optimized on data and MC. The uncertainties are statistical only.

As a summary, Tables 4.17, 4.18 and 4.19 show the different set of cuts which maximize the OS and SSπ performances for 35 pb^{-1}, 370 pb^{-1} and 1 fb^{-1} labeled as *Moriond 2011*, *LeptonPhoton 2011* and *Moriond 2012*.

To obtain a reliable mistag estimation to be used event-per-event in the CP fits η has to be calibrated. Finally, we can measure ω and monitor the performances of the taggers. Due to the differences of the offline and trigger selections, the p_T spectrum of the selected B events and the event multiplicity, small deviations in the tagging performances can be found between different control channels. For instance, the performance in the $B^0 \to D^-\pi^+$ channel is higher than in J/ψ channels where there are lower p_T cuts and a bigger fraction of TOS events in the trigger selections. In Tables 4.13, 4.14, 4.15 and 4.16 there is a summary of the tagging performances for the OS taggers with different control channels with 1 fb^{-1}.

Taggers	$\epsilon_{tag}[\%]$	ω [%]	$\epsilon_{tag}(1-2\omega)^2$ [%]
μ	5.20±0.04	30.8±0.4	0.77±0.04
e	2.46±0.03	30.9±0.6	0.36±0.07
K	17.67±0.08	39.33±0.24	0.81±0.04
$Q_{\rm vtx}$	18.46±0.08	40.31±0.24	0.70±0.04

TABLE 4.13: Tagging performance for 1 fb^{-1} in the $B^+ \to J/\psi K^+$ channel. Uncertainties are statistical.

Taggers	$\epsilon_{tag}[\%]$	ω [%]	$\epsilon_{tag}(1-2\omega)^2$ [%]
μ	5.18±0.07	33.8±1.0	0.54±0.07
e	2.49±0.05	36.1±1.4	0.19±0.04
K	18.0±0.1	40.0±0.6	0.72±0.09
$Q_{\rm vtx}$	18.4±0.1	39.9±0.6	0.75±0.09

TABLE 4.14: Tagging performance for 1 fb^{-1} in the $B^0 \to J/\psi K^*$ channel. Uncertainties are statistical only.

Taggers	ϵ_{tag}[%]	ω [%]	$\epsilon_{tag}(1-2\omega)^2$ [%]
μ	6.08±0.04	33.3±0.4	0.68±0.04
e	2.49±0.02	34.3±0.7	0.25±0.02
K	13.36±0.05	38.3±0.3	0.74±0.04
$Q_{\rm vtx}$	16.53±0.06	41.5±0.3	0.48±0.03

TABLE 4.15: Tagging performance for 1 fb^{-1} in the $B^0 \to D^{*-}\mu^+\nu_\mu$ channel. Uncertainties are statistical only.

Taggers	ϵ_{tag}[%]	ω [%]	$\epsilon_{tag}(1-2\omega)^2$ [%]
μ	7.47±0.07	30.0±0.7	1.19±0.8
e	3.09±0.05	32.7±1.1	0.37±0.5
K	14.70±0.10	37.9±0.5	0.86±0.7
$Q_{\rm vtx}$	14.23±0.10	37.1±0.5	0.95±0.8

TABLE 4.16: Tagging performance for 1 fb^{-1} in the $B^0 \to D^+\pi^-$ channel. Uncertainties are statistical only.

For the SSπ the performances with 1 fb^{-1} could be measured in $B^0 \to J/\psi K^*$ and $B^0 \to D^-\pi^+$ channels with an ϵ_{eff} of 0.75±0.08 % and 1.12±0.08 % respectively. In case of the SSK, the tagging performances were measured on 500 pb^{-1} in the $B_s^0 \to D_s^-\pi^+$ channel with a $\epsilon_{tag} = 16.3 \pm 0.4\%$, a wrong tag fraction $\omega = 35.3 \pm 2.1\%$ giving an $\epsilon_{\mathit{eff}} = 1.4 \pm 0.4\%$.

tagger	min p_T [GeV/c]	min p [GeV/c]	min IP/σ_{IP}	max track χ^2/ndf	Particle Identification Cuts	min $IP_{PU}/\sigma_{IP_{PU}}$	min $\lvert\Delta\phi\rvert$	extra cuts	$Prob_{min}$ $(1-\omega)$
Cuts optimized on 2010 data - Moriond 2011									
μ	1.1	-	-	2.2	$\Delta LL_{\mu-\pi} > 2$	3	5mrad	NSH CloneDistance	-
e	1	-	2	5	$\Delta LL_{e-\pi} > 4$	3	30mrad	VeloChMax <1.6 $E/p > 0.6$	-
OS K	0.8	4	4	2.7	$\Delta LL_{K-\pi} > 5$ $\Delta LL_{K-\pi} - \Delta LL_{p-\pi} > -4$	6	5mrad	$\lvert IP\rvert < 1.5$mm NSH	
OS K(*)	0.8	5.875	4.05	2.125	$\Delta LL_{K-\pi} > 6.5$ $\Delta LL_{K-\pi} - \Delta LL_{p-\pi} > -3.5$	4.72	5mrad	$\lvert IP\rvert < 1.25$mm NSH	0.51
Cuts optimized on summer 2011 data - LP 2011									
μ	1.2	-	-	3.2	$\Delta LL_{\mu-\pi} > 2.5$	3	5mrad	NSH CloneDistance	-
e	1	-	2	5	$\Delta LL_{e-\pi} > 4$	3	30mrad	VeloChMax <1.6 $E/p > 0.6$	-
OS K	0.8	5.875	4.05	2.125	$\Delta LL_{K-\pi} > 6.5$ $\Delta LL_{K-\pi} - \Delta LL_{p-\pi} > -3.5$	4.72	5mrad	$\lvert IP\rvert < 1.25$mm NSH	0.51
Cuts optimized on 2011 data - Moriond 2012									
μ	1.2	-	-	3.2	$\Delta LL_{\mu-\pi} > 2.5$	3	5mrad	NSH CloneDistance	-
e	1	-	1	3.85	$\Delta LL_{e-\pi} > 4$	5	5mrad	VeloChMax <1.6 $E/p > 0.75$	-
OS K	0.7	5	4.3	2.15	$\Delta LL_{K-\pi} > 0.75$ $\Delta LL_{K-\pi} - \Delta LL_{p-\pi} > -3$	7.5	5mrad	$\lvert IP\rvert < 1.45$mm NSH	0.54

TABLE 4.17: Summary of the optimized cuts for OS muon, electron, kaon taggers obtained with $B^0 \to D^{*-}\mu^+\nu_\mu$ and $B^+ \to J/\psi K^+$ events. This cuts are used to define individual tagger candidates presented in Sections 3.3 and 3.4 and are superseeded with the cuts in Table 3.1.

SS π	min p_T [GeV/c]	min p [GeV/c]	max IP/σ_{IP} max $IP(mm)$	max track χ^2/ndf	Particle Identification Cuts	$\Delta\eta$, $\Delta\phi$ and ΔR	ΔQ	$Prob_{min}$ $(1-\omega)$
Cuts optimized on 2010 data - Moriond 2011								
	0.75	5	3.5	5	$\Delta LL_{K-\pi} < 3$ $\Delta LL_{p-\pi} < 10$	- -	$\Delta Q < 1.5$GeV/c^2	0.53
Cuts optimized on summer 2011 data - LP 2011								
	0.75	5	3.5	5	$\Delta LL_{K-\pi} < 3$ $\Delta LL_{p-\pi} < 10$	- -	$\Delta Q < 1.5$GeV/c^2	0.54
Cuts optimized on 2011 data - Moriond 2012								
	0.6	4	9	5	$\Delta LL_{K-\pi} < 4.3$ $\Delta LL_{p-\pi} < 14$	-0.5$< \Delta\eta <$ 0.35 $\lvert\Delta\phi\rvert < 0.75$ $\lvert\Delta R\rvert < 0.75$	$\Delta Q < 1.2$GeV/c^2	0.56

TABLE 4.18: Summary of the optimized cuts for SSπ with $B^0 \to D^{*-}\mu^+\nu_\mu$ and $B^+ \to J/\psi K^+$ events. This cuts are used to define the individual tagger candidates presented in Sections 3.3 and 3.4 and are superseeded with the cuts in Table 3.1.

OS Vch	Seed Prob	Power K	min abs(Vch)	min $\Sigma_i p_T^i$ [GeV/c^2]	min $\Sigma_i IP^i/\sigma_{IP}^i$	min $\Sigma_i p^i$ [GeV/c^2]	min $\Sigma_{i,j}^{i\neq j} DOCA^{ij}$ [mm]	min $\Sigma_{i,j}^{i\neq j} m^{ij}$ [GeV/c^2]	$Prob_{min}$ $(1-\omega)$
Cuts optimized on 2010 data - Moriond 2011									
	0.4	0.4	0.17	-	-	-	-	-	0.53
Cuts optimized on summer 2011 data - LP 2011									
	0.4	0.4	0.275	1.5	10	10	0.5	0.5	0.54
Cuts optimized on 2011 data - Moriond 2012									
	0.42	0.55	0.2	1.55	10	8	0.5	0.6	0.54

TABLE 4.19: Summary of the optimized cuts for the OS Vertex charge tagger presented in Section 3.3.4 with $B^0 \to D^{*-}\mu^+\nu_\mu$ and $B^+ \to J/\psi K^+$ events. This cuts are used superseeded with the cuts in Table 3.1.

Chapter 5

Combination of flavour tagging algorithms

To be able to perform CP violation and B meson oscillations measurements, we need to combine multiple individual tagging decisions into a single decision. This chapter explains how the outputs of the different tagging algorithms can be combined and how the use of the per-event wrong tag fraction provides the best tagging power.

5.1 Introduction

In each tagger i, the charge of the tagging particle (or vertex charge) is used to define a tagging decision d_i which indicates the initial flavour of the reconstructed signal B candidate ($d = -1$ when the signal B hadron contains a b quark and $d = +1$ when it contains a $\bar{b}$ quark). When more than one tagger is available per event, different strategies[1] can be used to combine their decisions into a global tagging decision.

5.2 Combination of taggers

The flavour tagging as a global algorithm assigns an individual probability $p_i = 1 - \omega_i$ of being correct to each separate tagger i. This probability is based on the estimation of the mistag fraction η_i which is a function of the event properties and of the kinematic and geometrical properties of the taggers themselves. As the mistag estimation is evaluated by means of a neural net based on Monte Carlo, a calibration on data must be performed to properly estimate

[1] Available in the Flavour Tagging (FT) and Flavour Tagging Checker (FTC) packages

the mistag fraction. Once the tagging decision of each tagger is optimized and calibrated individually, these five separate decisions (one for each tagger) can be combined to get a single decision. The combined probability $\mathcal{P}(b)$ of the B meson to contain a b quark is calculated as:

$$\mathcal{P}(b) = \frac{p(b)}{p(b) + p(\bar{b})}, \qquad \mathcal{P}(\bar{b}) = 1 - \mathcal{P}(b), \tag{5.1}$$

where $p(b) = \prod_i p_i(b)$, being $p_i(b)$ the probability to have a b-tagged meson as a response from the tagger i. If the tagging decision is -1, then $p_i(b) = p_i$ while if the tagging decision is $+1$, $p_i(b) = 1 - p_i$. This can be summarized in a more compact way by the following equations:

$$p(b) = \prod_i (\frac{1 + d_i}{2} - d_i p_i), \qquad p(\bar{b}) = \prod_i (\frac{1 - d_i}{2} + d_i p_i). \tag{5.2}$$

where $p_i = 1 - \eta_i$, η_i is the estimated mistag fraction associated to the i-th tagger, and $d_i = \pm 1$ is the decision taken by the i-th tagger. The final combined decision is taken comparing the values of $\mathcal{P}(b)$ and $\mathcal{P}(\bar{b})$. If $\mathcal{P}(b) > \mathcal{P}(\bar{b})$ the combined tagging decision is $d = -1$ with a mistag fraction $\omega = 1 - \mathcal{P}(b)$, while if $\mathcal{P}(\bar{b}) > \mathcal{P}(b)$ the tagging decision is $d = 1$ and $\omega = 1 - \mathcal{P}(\bar{b}) = \mathcal{P}(b)$. As in the individual taggers, a cut on the probability $\mathcal{P}(b)$ can also be applied to exclude events with ω close to 0.5.

Two possible combinations are available, using only OS taggers, and combining OS and SS taggers. In particular, the combination of SS kaon tagger with OS taggers can increase the ϵ_{eff} by almost a factor 2 [87]. This can be seen in Tables 5.1 and 5.2, where the inclusion of the SS kaon tagger increases the combined effective tagging efficiency of MC simulated events, from 3.40% to 6.46% in the case of $B_s^0 \to J/\psi\phi$, and from 4.53% to 8.67% in the case of the $B_s^0 \to D_s^- \pi^+$ mode. In these two tables the performances are also split into *tagging categories*. Their meaning and definition are described in Section 5.4. In the case of the SSπ, a relative increase of $\sim 50\%$ can be achieved [87].

Due to the correlation among taggers, which is neglected in Eq. 5.2, the combined mistag fraction results slightly underestimated, as can be seen in Figure 5.1. The main source of correlation is between the secondary vertex charge tagger and the other opposite side taggers, since sometimes the secondary vertex can be build with the same particles used by other opposite side taggers. In Table 5.3 we summarize the correlations between the single taggers evaluated from MC $B^+ \to J/\psi K^+$ events (and crosschecked on MC $B_s^0 \to J/\psi\phi$ events) and the ones evaluated from 1 fb^{-1} of $B^+ \to J/\psi K^+$ events in the signal region from the 2011 LHCb physics run. The correlations on both samples are compatible. From the sum of the relative efficiencies

$B_s^0 \to J/\psi\phi$			
	ϵ_{tag} (%)	ω (%)	ϵ_{eff} (%)
Individual taggers			
μ	6.02±0.08	31.1±0.6	0.86±0.06
e	2.91±0.06	31.0±0.9	0.42±0.04
OS K	15.60±0.12	35.3±0.4	1.34±0.07
SS K	26.75±0.15	34.9±0.3	2.43±0.10
Q_{vtx}	34.64±0.16	40.9±0.3	1.14±0.07
Combination of OS taggers only			
cat 1	27.18±0.15	44.9±0.3	0.28±0.04
cat 2	6.97±0.08	35.5±0.6	0.59±0.05
cat 3	5.02±0.07	30.8±0.7	0.74±0.05
cat 4	3.68±0.06	26.0±0.8	0.84±0.06
cat 5	2.13±0.05	16.7±0.8	0.95±0.05
average	45.0±0.2	39.0±0.2	2.19±0.10
combined	45.0±0.2	36.2±0.2	3.40±0.11
Combination of all taggers			
cat 1	27.46±0.15	42.8±0.3	0.57±0.05
cat 2	11.06±0.10	35.1±0.5	0.98±0.06
cat 3	7.61±0.09	28.9±0.5	1.35±0.07
cat 4	5.14±0.07	23.2±0.6	1.48±0.07
cat 5	4.20±0.07	14.8±0.6	2.08±0.08
average	55.47±0.16	35.8±0.2	4.50±0.13
combined	55.47±0.16	32.93±0.2	6.46±0.15

TABLE 5.1: Flavour tagging performance for offline selected MC $B_s^0 \to J/\psi\phi$ events passing the Level-0 trigger, for the individual taggers and for their combination. Results using opposite side tagger only are shown as well. Uncertainties are statistical.

on the diagonal, an overall correlation of $\sim 30\%$ and $\sim 23\%$ is calculated in MC and data respectively. The correlation between SS and OS taggers is very small ($< 1\%$ [87]) and can be neglected.

To correct for the underestimation of η_{OS} due to correlation effects between the taggers, the FT applies a correction of the combined OS tagging decision and mistag estimation. This correction corresponds to the calibration on data described in Section 5.3.

Once the combined OS tagging decision is obtained, it can be combined with the SS decision according to Eq. 5.1 and 5.2 to obtain a final decision. A further correction is not needed anymore as the OS and SS taggers are not correlated.

5.3 Calibration of the OS mistag

As explained above, due to the correlations among taggers neglected in Eq. 5.2, the combined mistag fraction is slightly underestimated. As discussed in Chapter 4, a linear dependence between the estimated η and the measured ω for signal events is used:

$B_s^0 \to D_s^- \pi^+$			
	ϵ_{tag} (%)	ω (%)	ϵ_{eff} (%)
Individual taggers			
μ	9.2±0.1	29.0±0.6	1.6±0.1
e	2.9±0.1	30.6±1.1	0.43±0.05
OS K	15.0±0.1	31.6±0.5	2.0±0.1
SS K	26.0±0.2	30.9±0.4	3.8±0.2
$Q_{\rm vtx}$	39.5±0.2	40.0±0.3	1.6±0.1
Combination of OS taggers only			
cat 1	27.88±0.18	44.2±0.4	0.38±0.05
cat 2	7.93±0.11	34.4±0.7	0.77±0.07
cat 3	5.95±0.09	28.9±0.7	1.06±0.07
cat 4	4.33±0.08	25.0±0.8	1.09±0.07
cat 5	3.11±0.07	18.4±0.9	1.24±0.07
average	49.2±0.2	37.4±0.3	3.10±0.14
combined	49.2±0.2	34.8±0.3	4.53±0.15
Combination of all taggers			
cat 1	27.81±0.18	43.1±0.4	0.53±0.06
cat 2	10.77±0.12	33.3±0.6	1.21±0.08
cat 3	8.22±0.11	27.3±0.6	1.70±0.10
cat 4	6.18±0.10	23.7±0.7	1.70±0.09
cat 5	6.83±0.10	14.0±0.5	3.54±0.12
average	59.8±0.2	33.8±0.2	6.25±0.19
combined	59.8±0.2	31.0±0.2	8.67±0.20

TABLE 5.2: Flavour tagging performance for offline selected MC $B_s^0 \to D_s^- \pi^+$ events passing the Level-0 trigger, for the individual taggers and for their combination. Uncertainties are statistical.

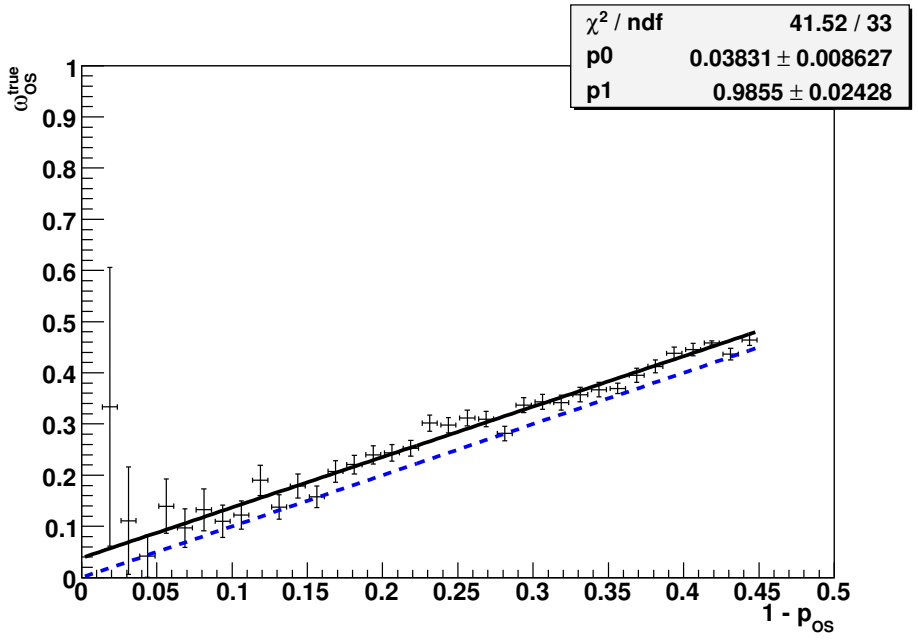


FIGURE 5.1: True wrong tag fraction of the combined OS taggers versus the estimated probability of a wrong tag obtained as a product of individual OS taggers probabilities. For MC $B_s^0 \to J/\psi\phi$ offline selected events after the L0 trigger. The solid line corresponds to a linear fit and the dashed line to the expected behavior in the absence of correlations.

tagger	μ	e	K	Q_{vtx}
μ	18%	0.3%	2.6%	6.3 %
e	–	8.7%	1.6%	3.3 %
K	–	–	44 %	18.3%
Q_{vtx}	–	–	–	59 %

tagger	μ	e	K	Q_{vtx}
μ	19%	0.34%	2.9%	5.1 %
e	–	7.8%	1.6%	2.2 %
K	–	–	42 %	13%
Q_{vtx}	–	–	–	54 %

TABLE 5.3: Correlation among taggers for OS tagged events calculated on MC (left) and ~ 1 fb^{-1} of data (right) in $B^+ \to J/\psi K^+$ signal events. The diagonal terms are the relative efficiencies (in %) of the single taggers. The off-diagonal terms are the relative efficiencies of combination of two or more taggers. From the sum of the diagonal terms an overall correlation of $\sim$ 30% and $\sim$ 23% are calculated for MC and data respectively. It is mainly due to a correlation of the single particle taggers with the Q_{vtx} one.

$$\omega = p_0 + p_1 \cdot (\eta - <\eta>) \tag{5.3}$$

where p_0 and p_1 are free parameters and $<\eta>$ is the mean mistag probability. This parametrization is chosen to minimize the correlation between the two parameters. Any deviation from $p_0 =<\eta>$ and $p_1 = 1$ indicates that the calculated mistag should be corrected. In order to extract the calibration parameters an unbinned maximum likelihood fit was performed on signal events of the self-tagging $B^+ \to J/\psi K^+$ control channel as explained in Chapter 4. The calibration was performed twice, first the sample was split into two random samples, and since the values obtained in the two independent samples where statistically compatible with the other sample, the calibration was repeated using the whole data sample to reduce the statistical uncertainty.

After the calibration, the values of p_0, p_1 and $<\eta>$ were re-measured in the whole $B^+ \to J/\psi K^+$ sample. As expected, the ω distribution is flat for background events while for signal events the p_0 and p_1 parameters are compatible with $<\eta>$ and 1 ensuring that the $B^+ \to J/\psi K^+$ sample is properly calibrated. In Figure 5.2 both distributions are shown for the integrated luminosity of 370 pb^{-1}.

To make sure that the calibration can be exported to other channels, the values of p_0, p_1 and $<\eta>$ were crosschecked on the $B^0 \to D^{*-}\mu^+\nu_\mu$ and $B^0 \to J/\psi K^*$ control channels. In Table 5.4 the results for the calibration parameters with 370 pb^{-1} of data for the $B^+ \to J/\psi K^+$, $B^0 \to D^{*-}\mu^+\nu_\mu$ and $B^0 \to J/\psi K^*$ control channels are reported after the calibration procedure was applied.

The $B^0 \to J/\psi K^*$ channel is compatible with the calibration done in the $B^+ \to J/\psi K^+$ channel, but with larger statistical uncertainty. In the $B^0 \to D^{*-}\mu^+\nu_\mu$ channel the trigger and offline selections as well as the signal spectra are quite different from the J/ψ channels. Therefore, the agreement in the resulting calibration parameters can be regarded as an additional validation of the calibration and its universality. In Figure 5.3 the measured OS mistag fraction as a function of the estimated mistag probability is shown for $B^0 \to D^{*-}\mu^+\nu_\mu$ signal events after

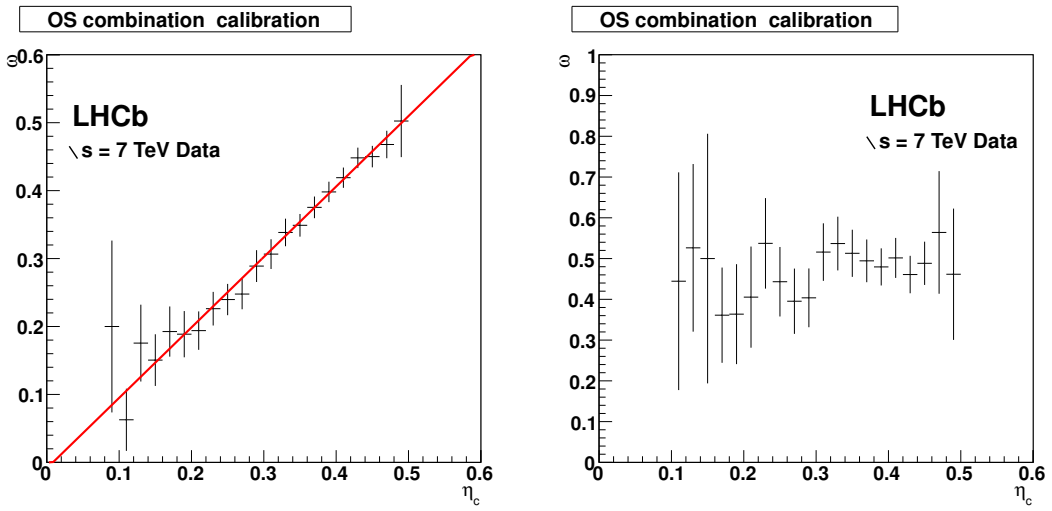


FIGURE 5.2: Calibration plots for the signal (left) and the background (right) from the fit to the $B^+ \to J/\psi K^+$ channel after applying the calibration. Only OS combination and 370 pb^{-1} were used. Black points are data, the solid red line is the fit function.

channel	p_0	p_1	$< \eta_c >$	$p_0 - p_1 < \eta_c >$	$\rho(p_0, p_1)$
$B^+ \to J/\psi K^+$	0.384 ± 0.003	1.037 ± 0.038	0.379	-0.009 ± 0.014	0.14
$B^0 \to J/\psi K^*$	0.399 ± 0.008	1.016 ± 0.102	0.378	0.015 ± 0.039	0.05
$B^0 \to D^{*-}\mu^+\nu_\mu$	0.396 ± 0.002	1.023 ± 0.026	0.375	0.008 ± 0.010	0.14

TABLE 5.4: Fit values and correlations of the mistag calibration parameters measured in the $B^+ \to J/\psi K^+$, $B^0 \to J/\psi K^*$ and $B^0 \to D^{*-}\mu^+\nu_\mu$ channels for the OS combination on 370 pb^{-1}. The uncertainty is statistic.

the calibration. The linear fit superimposed corresponds to the parametrization of Eq. 5.3 and the parameters of Table 5.4.

The same calibration process was performed with 1 fb^{-1} and analogous results were found. Table 5.5 shows the results on the calibration parameters after the calibration procedure was applied. In Figure 5.4 the calibration for $B^0 \to J/\psi K^*$, $B^0 \to D^-\pi^+$ and $B^0 \to K^+\pi^-$ channels is shown.

channel	p_0	p_1	$< \eta_c >$	$\rho(p_0, p_1)$
$B^+ \to J/\psi K^+$	0.392 ± 0.002	1.035 ± 0.021	0.391	0.12
$B^0 \to J/\psi K^*$	0.400 ± 0.004	1.013 ± 0.053	0.390	0.06
$B^0 \to D^-\pi^+$	0.398 ± 0.003	1.010 ± 0.039	0.393	
$B^0 \to K^+\pi^-$	0.355 ± 0.014	0.99 ± 0.16	0.353	0.14

TABLE 5.5: Fit values and correlations of the mistag calibration parameters measured in the $B^+ \to J/\psi K^+$, $B^0 \to J/\psi K^*$, $B^0 \to D^-\pi^+$ and $B^0 \to K^+\pi^-$ channels for the OS combination with 1 fb^{-1}. The uncertainty is statistic.

Once the OS tagger mistag probability is calibrated, it can be combined with the calibrated SS tagger mistag probability. With 370 pb^{-1} only the SSπ was calibrated. In Figure 5.5, the

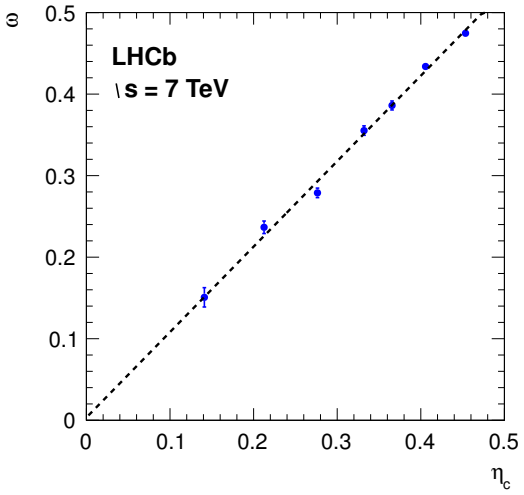


FIGURE 5.3: Distribution of the measured OS mistag fraction (ω) versus calibrated mistag probability (η_c), for background subtracted $B^0 \to D^{*-}\mu^+\nu_\mu$ events in the signal mass window. Points with errors are data, the lines represent the result of the fit corresponding to the parameters in Table 5.4.

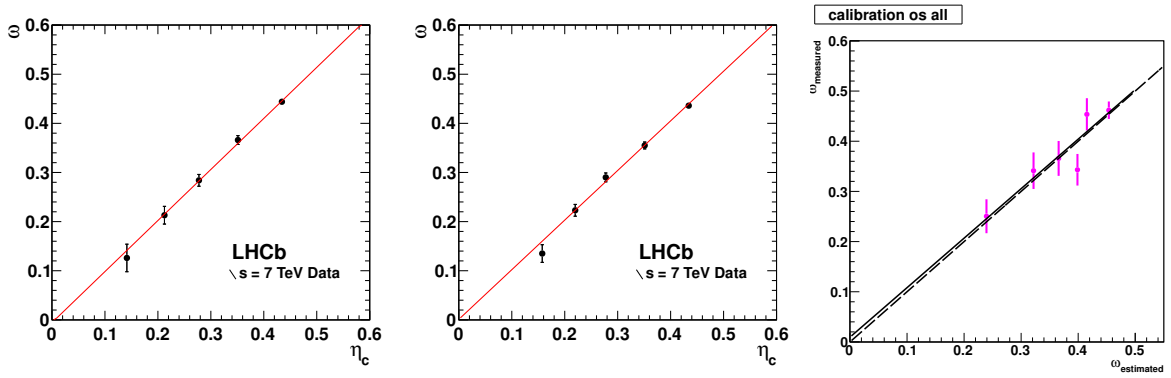


FIGURE 5.4: From left to right: Calibration plot for the OS tagger combination determined for the fit at $B^0 \to J/\psi K^*$, $B^0 \to D^-\pi^+$ and $B^0 \to K^+\pi^-$ channels with 1 fb^{-1}. Black points are data, the solid line is the fit function.

calibration plots for the SSπ and OS combination are shown, as expected, the calibration values are already correct due to the fact that the SS and the OS are not correlated. With 1 fb^{-1} the calibration of the SSπ and SSK is being performed with B^0 and B_s^0 channels respectively.

Other alternatives to obtain an uncorrelated combined OS tagging decision were envisaged in the past. However, the calibration of the OS η from data was the method finally used. In the following sections such methods are explained in detail and their performances are tested on MC data giving compatible or in some cases even better results than the default combination which uses Eq. 5.1 and Eq. 5.2.

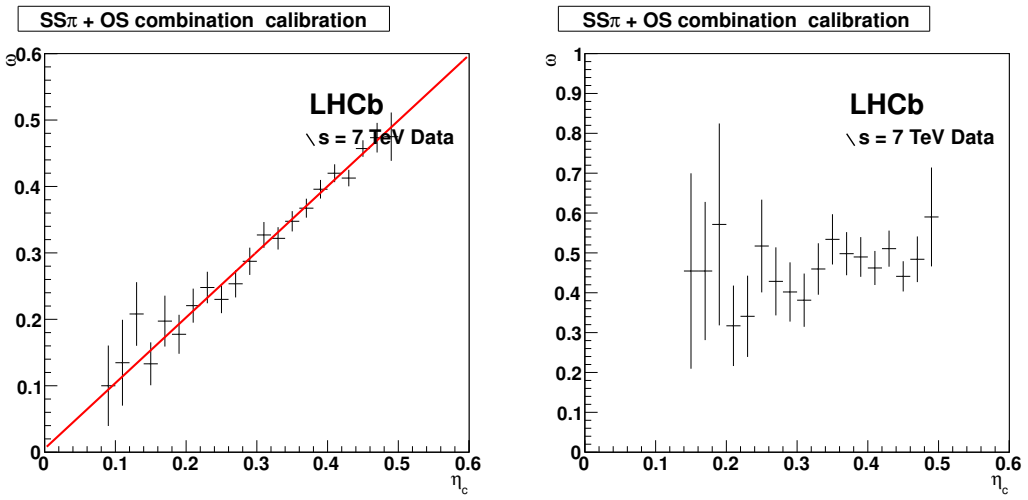


FIGURE 5.5: Calibration plots for the signal (left) and the background (right) from the fit to the $B^+ \to J/\psi K^+$ channel. SS and OS combination is shown. 370 pb^{-1} were used. Black points are data, the solid red line is the fit function.

5.3.1 Excluding the Secondary Vertex Charge Tagger

As the main correlation in the OS combination is between the Secondary Vertex Charge (SVC) tagger and other OS taggers, we tried to exclude the vertex tagger. The effect of removing the SVC tagger from the combination of taggers is shown in Figure 5.6, which has to be compared with Figure 5.1, where the tagger is activated. When the vertex charge is excluded from the OS combination the underestimation of ω disappears as the main correlation between the taggers is gone. The total effective efficiency in this case is reduced from 6.46% to 6.05% for MC $B_s^0 \to J/\psi\phi$ events.

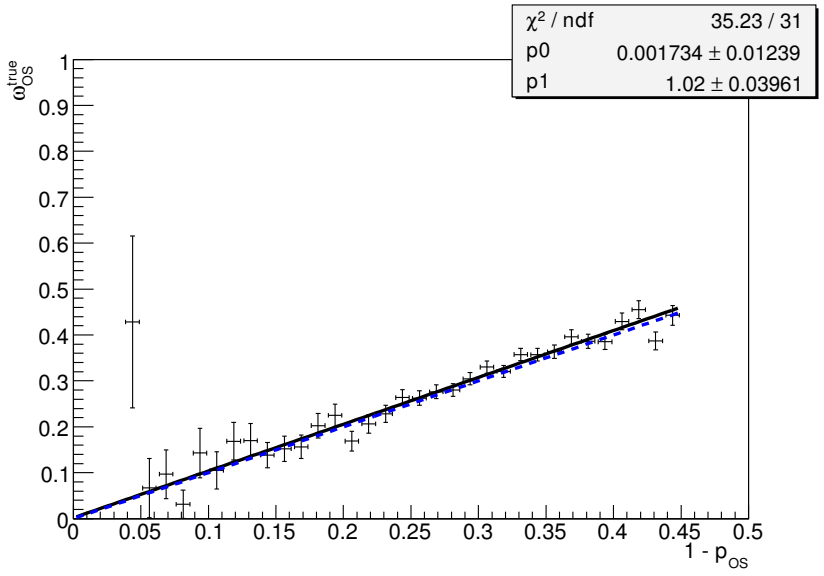


FIGURE 5.6: True wrong tag fraction of the combined OS taggers versus the estimated probability to be wrong excluding the OS vertex charge tagger.

If we remove the SVC tagger when one of the other taggers is present, the effective efficiency decreases to 6.26%. Since the SS tagger response shows a negligible correlation with the SVC tagger, we can use SVC tagger also when SS tagger is present. The effective efficiency in this case goes from 6.26% to 6.33% in MC $B_s^0 \to J/\psi\phi$ events. This can be done without introducing a significant bias in the η estimation as seen from Figure 5.6. Table 5.6 shows the tagging performance obtained in this last case, that has to be compared with Table 5.1. This option has the advantage of being very simple, with a modest reduction of the tagging performance.

$B_s^0 \to J/\psi\phi$			
	ϵ_{tag} (%)	ω (%)	ϵ_{eff} (%)
cat 1	26.98±0.15	42.7±0.3	0.57±0.05
cat 2	12.19±0.11	34.3±0.4	1.21±0.07
cat 3	7.34 ±0.09	26.8±0.5	1.57±0.08
cat 4	3.95 ±0.06	21.0±0.7	1.32±0.07
cat 5	3.17 ±0.06	13.9±0.6	1.65±0.07
average	53.6±0.2	34.1±0.2	4.93±0.15
combined	53.6±0.2	32.8±0.2	6.33±0.15

TABLE 5.6: Tagging performance in $B_s^0 \to J/\psi\phi$ decays after the L0, using $Q_{\rm vtx}$ as long as it is the only available tagger or in presence of the SS tagger. It has to be compared with Table 5.1.

5.3.2 Use of a Neural Net for the combination of taggers

One valuable feature of Neural Nets is that they are sensitive to correlations in the input variables, so that if one of the tagger decisions was an exact duplicate of another one (i.e. 100% correlated), its weight would trend automatically to zero by the Neural Net in the training phase.

The simple idea is therefore to feed the Neural Net inputs with the output ω_i from the 5 individual tagger probability estimation. In the case that the tagger decision is $\bar{B}$, then $1 - \omega_i$ is given as an input.

Treating B-tagged events as "signal"(1) and $\bar{B}$-tagged events as "background"(0) is a technical convenient way to exploit one of the multi-variate tools like MLP [82]. Events which are tagged by one single tagger do not obviously need any special treatment, so the Neural Net is only trained and applied on events with multiple active taggers. As SS tagging is different for $B_{d,u}$ and B_s channels, two different trainings for the Neural Net combination would be needed in theory. As a matter of fact this turns out to be unnecessary, as applying the tuning of a B_s channel to a B_d channel shows the same tagging performance that one obtained with a dedicated tuning. This is not so surprising if one considers that in the case of the OS taggers, the correlations are similar in the B_s and $B_{d,u}$ modes, while for the OS versus SS they are anyway small. As a result, one single Neural Net is able to deal with both cases at once.

The results with this method are better than the ones of the previous method, giving a ϵ_{eff}=6.55 $\pm$ 0.15% on MC $B_s^0 \to J/\psi\phi$ events (see Table 5.7), and ϵ_{eff}=8.80 $\pm$ 0.21% for MC $B_s^0 \to D_s^-\pi^+$. Other channels have been studied without changing the tuning of the neural net itself, obtaining similar or better tagging performances and successfully correcting the correlation issues. Figure 5.7 shows the true mistag as a function of the estimated mistag for the three channels $B_s^0 \to J/\psi\phi$, $B_s^0 \to D_s^-\pi^+$ and $B^+ \to J/\psi K^+$on MC data. This procedure allows to have an event-per-event correct estimation of ω which can then be directly used as input to the CP fits.

$B_s^0 \to J/\psi\phi$			
	ϵ_{tag} (%)	ω (%)	ϵ_{eff} (%)
cat 1	21.41±0.14	44.0±0.4	0.30±0.04
cat 2	13.29±0.11	37.2±0.4	0.87±0.06
cat 3	9.69 ±0.10	30.1±0.5	1.54±0.08
cat 4	6.24 ±0.08	22.9±0.6	1.83±0.08
cat 5	4.01 ±0.06	14.5±0.6	2.02±0.07
average	54.63±0.2	34.2±0.2	5.17±0.15
combined	54.63±0.2	32.7±0.2	6.55±0.15

TABLE 5.7: Performance in $B_s^0 \to J/\psi\phi$ decays after the L0, obtained combining the taggers using a dedicated neural net. To be compared with Table 5.1.

This method was applied on the OS and SS taggers combinations. As the correlation between SS and OS taggers is negligible, this method should be reapplied only for the OS taggers, where the correlation is dominant. The SS and OS can be combined independently as in Eq. 5.1 and 5.2 later on.

One additional possibility that has been explored is the application of a linear transformation to the η_i of the taggers defined by $\omega_i' = (1-k_i)\eta_i + k_i/2$ where k_i are 5 free parameters that minimize $|\omega_{\text{true}} - \omega_{\text{estimated}}|$. The minimum is found for $k_e = k_\mu = k_k = 0$, $k_{kSS} = 0.1$ and $k_{vtx} = 0.3$. The tagging performances are similar to those in the previous case, but the correction of the bias is not as good as in the case of the combination with the neural net.

5.4 Tagging categories

The use of tagging categories, albeit their small reduction of tagging power with respect to the event per event mistag, was initially considered as the default option for all the physics measurements that needed tagging, mainly because of its simplicity. And for this reason the use of tagging categories was extensively studied. In this Section we want to summarize the tagging methods that were developed.

The combined mistag probability can be used to categorize events according to their tagging purity in order to improve the global performances of the tagging. In this case, the events are

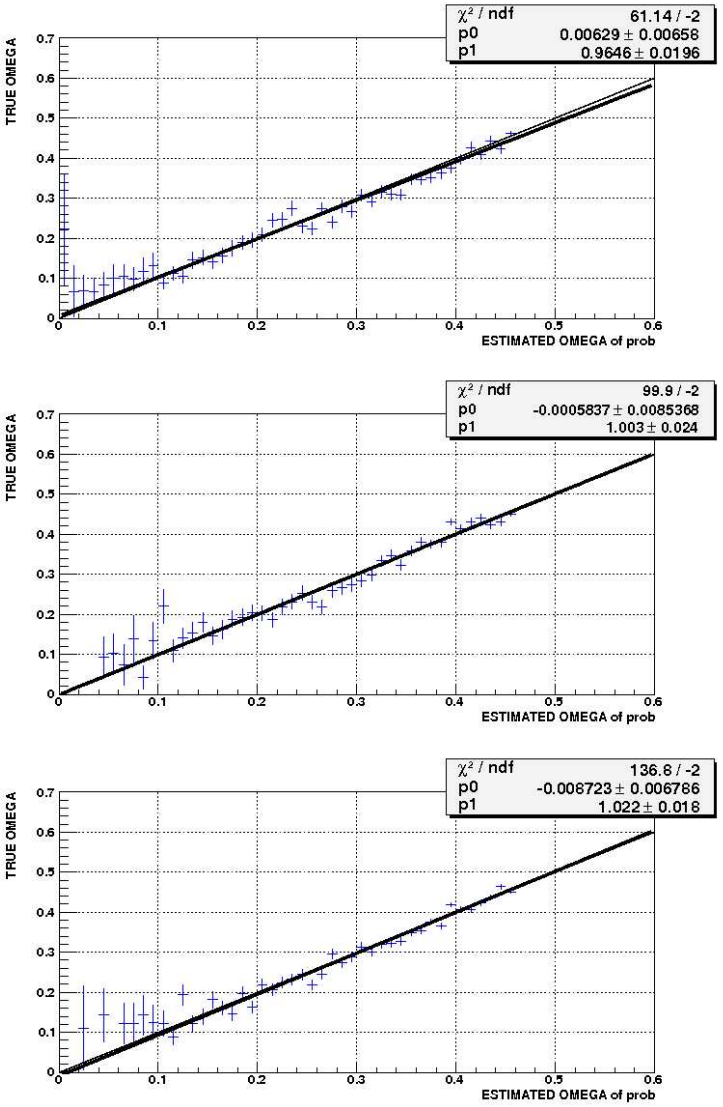


FIGURE 5.7: True ω versus the estimated wrong tag fraction from the Neural Net combination for the three channels $B_s^0 \to D_s^- \pi^+$ (top), $B_d \to \pi\pi$ (center) and $B^+ \to J/\psi K^+$ (bottom).

subdivided into a number of tagging categories of lowering ω. The splitting is obtained from the linear combination of Eq. 5.1. Alternatively, one possibility is to sort the events based on the type of taggers which contribute to form the tag decision (PID approach) as described in Section 5.4.1. The FT does not include any cut on the $\mathcal{P}(b)$ for the OS, and OS + SS combination, although for some analysis a cut could be introduced to improve the effective efficiency.

Figure 5.8 shows the outcome of the mistag probability distribution for each tagger and for the OS combination from Eq. 5.2. If we did not consider how events distribute in ω, but use them as such to evaluate CP asymmetries, then "bad" events (i.e. high ω) would mix up with "good" events (i.e. low ω) yielding the effective tagging efficiency that corresponds to the "Average" line in the Tables.

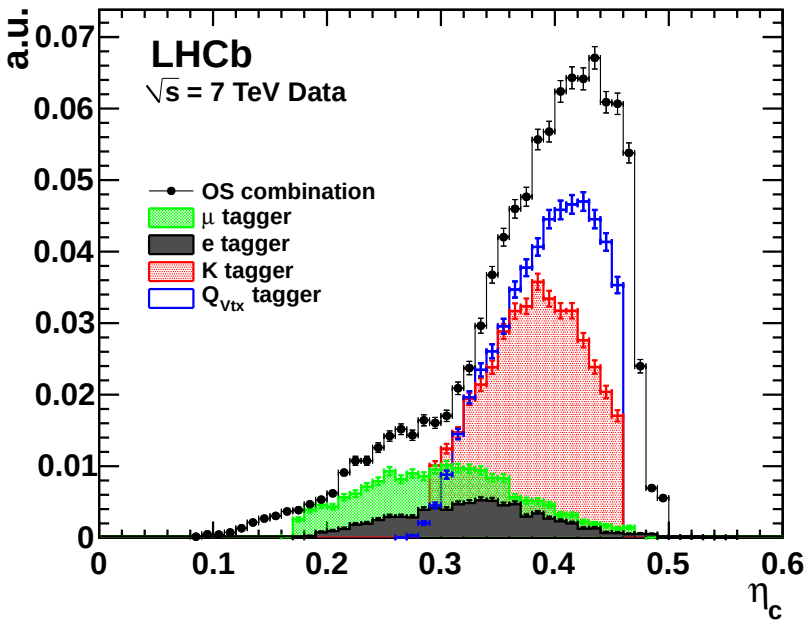


FIGURE 5.8: Distribution of the mistag probability for the $B^+ \to J/\psi K^+$channel signal events for the single OS taggers and their combination after calibration.

Grouping together events with similar ω, estimated from the neural net output, produces a significative improvement in the global performance of tagging. In this case the performance is determined in each independent category separately. The sum of the values determined in each of these categories separately will give us the total effective efficiency. This improvement can be seen from the "Combine" lines in the performance Tables.

In MC simulations, the events were sorted into 5 independent categories of decreasing mistag probability and a cut on $\mathcal{P}(b) < 0.56$ was applied to maximize the global tagging performance shown in the Tables 5.1 and 5.2.

The tagging performances for $B^0 \to D^{*-}\mu^+\nu_\mu$, $B^+ \to J/\psi K^+$ and $B^0 \to J/\psi K^*$ control channels splitted in tagging categories of increasing tagging purity are summarized in Tables 5.8, 5.9 and 5.10 for 370 pb^{-1} of data from the 2011 LHCb physics run. For the $B^0 \to D^{*-}\mu^+\nu_\mu$ and $B^0 \to J/\psi K^*$ channels ω was measured from a fit to the oscillation on independent samples as detailed in App. A. On Figure 5.9 the mixing asymmetry as a function of propertime is shown for $B^0 \to D^{*-}\mu^+\nu_\mu$ signal events for the 7 tagging categories.

Similar performances are obtained with 1 fb^{-1} of data from the whole 2011 physics run. They are detailed in Chapter 7.

$B^0 \to D^{*-}\mu^+\nu_\mu$ channel				
	ϵ %	ω %	$\epsilon(1-2\omega)^2$ %	–
Combination of taggers: OS Average values (η_c <0.42)				
OS	20.6±0.1	36.1±0.2	1.58±0.06	
Combination of taggers: OS Tagging categories (η_c <0.5)				
Categories	ϵ %	ω %	$\epsilon(1-2\omega)^2$ %	$< \eta_c >$
0.43< η_c	8.01±0.04	47.0±0.4	0.03±0.01	0.453
0.38< η_c ≤0.43	8.93±0.04	43.4±0.4	0.18±0.02	0.406
0.35< η_c ≤0.38	4.35±0.03	39.5±0.5	0.19±0.02	0.366
0.31< η_c ≤0.35	3.57±0.03	35.5±0.6	0.30±0.02	0.332
0.24< η_c ≤0.31	3.31±0.03	28.6±0.6	0.61±0.03	0.276
0.17< η_c ≤0.24	1.71±0.02	24.1±0.8	0.46±0.03	0.213
η_c ≤0.17	0.56±0.01	14.2±1.2	0.29±0.02	0.141
sum	30.48±0.08	37.0±0.3	2.05±0.06	

TABLE 5.8: Tagging performance in the $B^0 \to D^{*-}\mu^+\nu_\mu$ channel. Errors are statistical. The ω was obtained from the fit to the oscillation (see Figure 5.9). Average values for OS are obtained with a cut η_c <0.42.

$B^+ \to J/\psi K^+$channel				
	ϵ %	ω %	$\epsilon(1-2\omega)^2$ %	–
Combination of taggers: OS Average values (η_c <0.42)				
OS	17.81±0.13	34.6±0.4	1.69±0.10	
Combination of taggers: OS Tagging categories (η_c <0.5)				
Categories	ϵ %	ω %	$\epsilon(1-2\omega)^2$ %	$< \eta_c >$
0.38< η_c	15.92± 0.13	43.7±0.4	0.25±0.04	0.429
0.31< η_c ≤0.38	6.73±0.09	35.5±0.6	0.57±0.06	0.351
0.24< η_c ≤0.31	2.85±0.06	26.4±0.9	0.63±0.06	0.277
0.17< η_c ≤0.24	1.37±0.04	21.6±1.1	0.44±0.05	0.212
η_c ≤0.17	0.41±0.02	16.8±2.0	0.18±0.03	0.144
sum	27.3±0.2	36.2±0.5	2.07±0.11	

TABLE 5.9: Tagging performance in the $B^+ \to J/\psi K^+$ channel. Errors are statistical. Average values for OS are obtained with a cut η_c <0.42.

$B^0 \to J/\psi K^*$channel				
	ϵ %	ω %	$\epsilon(1-2\omega)^2$ %	–
Combination of taggers: OS Average values (η_c <0.42)				
OS	17.8±0.2	37.1±0.9	1.18±0.18	
Combination of taggers: OS Tagging categories (η_c <0.5)				
Categories	ϵ %	ω %	$\epsilon(1-2\omega)^2$ %	$< \eta_c >$
0.38< η_c	15.1±0.2	44.5±1.0	0.18±0.07	0.428
0.31< η_c ≤0.38	6.7±0.1	38.0±1.5	0.39±0.10	0.350
0.24< η_c ≤0.31	3.0±0.1	31.6±2.2	0.40±0.11	0.278
0.17< η_c ≤0.24	1.3±0.1	29.8±3.2	0.21±0.08	0.213
η_c ≤0.17	0.3±0.1	13.1±5.3	0.19±0.07	0.139
sum	26.5±0.3	38.6±0.8	1.38±0.20	

TABLE 5.10: Tagging performance in the $B^0 \to J/\psi K^*$ channel. Errors are statistical. Average values for OS are obtained with a cut η_c <0.42.

5.4.1 PID tagger type approach

One possibility to categorize events is to sort them by the tagger type, looking at the PID of the particle used by the taggers, according to the number of active taggers that give an answer (in each event). Independently of the probability of a tagger to be correct, we consider only events where the sum of the tagging decisions is different from zero.

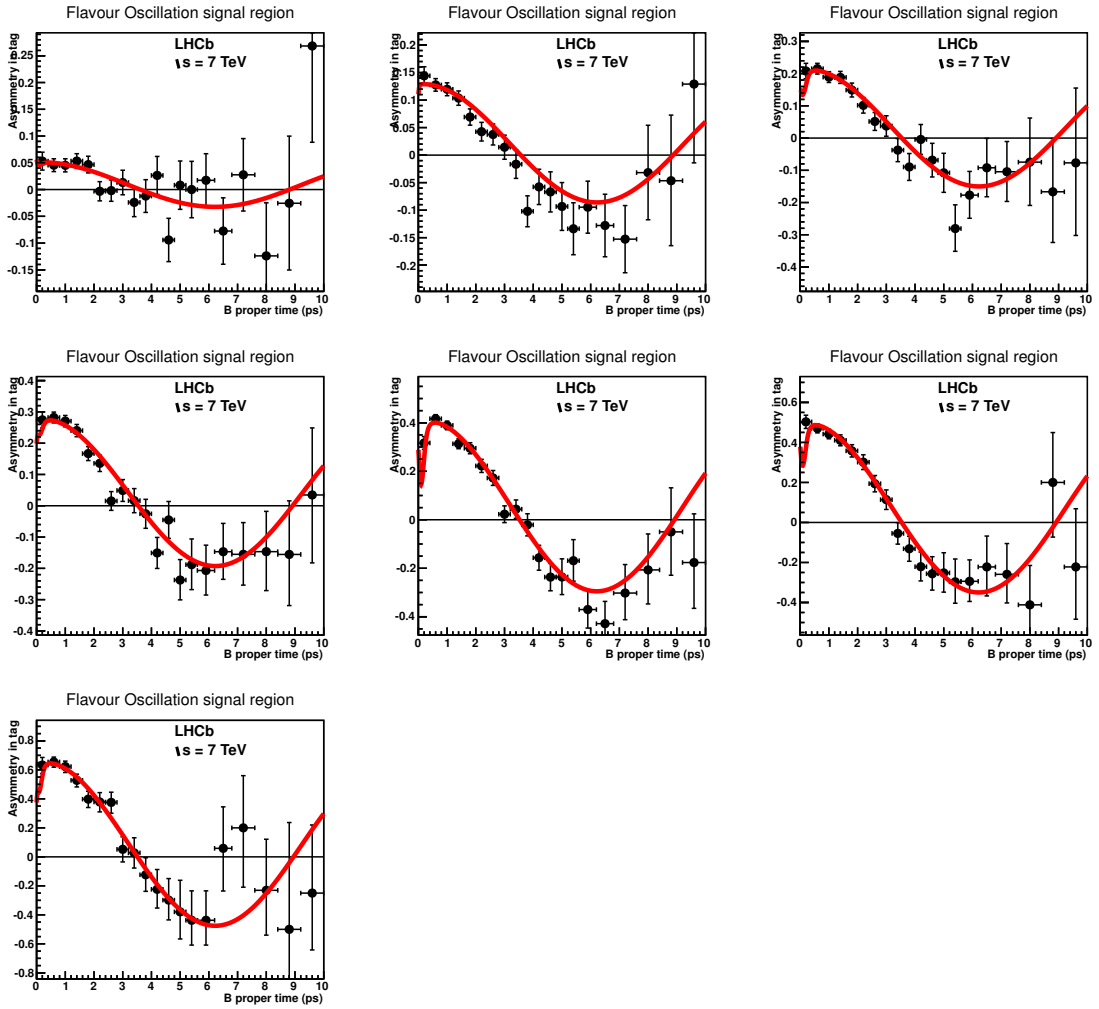


FIGURE 5.9: Mixing asymmetry as a function of proper-time in $B^0 \to D^{*-}\mu^+\nu_\mu$ events, in the signal mass region, using the OS tagger. Events are splitted into seven samples of increasing tagging purity using the mistag probability, from top left to bottom right categories from 1 to 7 [88].

The PID combination gives to each combination of active taggers a binary index $(nmlkji)$. Each digit corresponds to one tagger: muon (i), electron (j), OS kaon (k), SS pion or kaon (l) and vertex charge tagger (m). The (n) digit corresponds to an extra variable: the absolute value of the sum of the tagging decisions. If the absolute value of the sum of the tagging decision is greater than 1, meaning that there are at least two taggers which give the same answer, the digit would be 1 instead of zero. This is done because events where the absolute value of the sum of the tagging decisions is greater than 1 have an expected lower value of ω as shown in Figure 5.10.

For each combination of active taggers (for each index), we can compute the wrong tag fraction looking at the MC information or at the reconstructed B in a control channel. Then, the indeces are sorted by decreasing mistag fraction and are binned into 5 tagging categories of similar ω

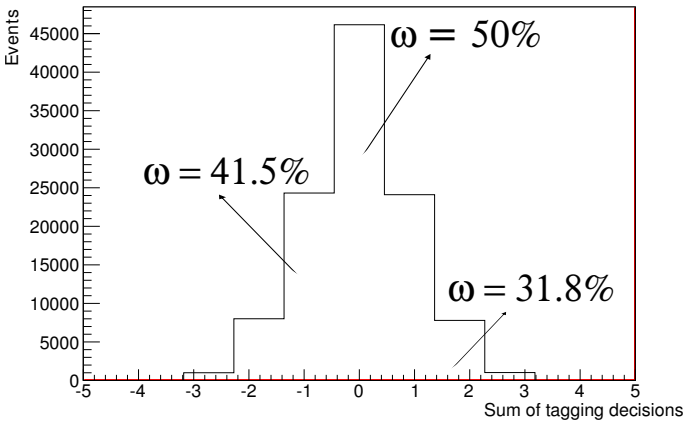


FIGURE 5.10: Sum of the tagging decision and their respective wrong tag fraction.

as in Figure 5.11. For each category, the average mistag is computed as well. The results with this method can be seen in Tables 5.11 (to be compared with Tables 5.7 and 5.1).

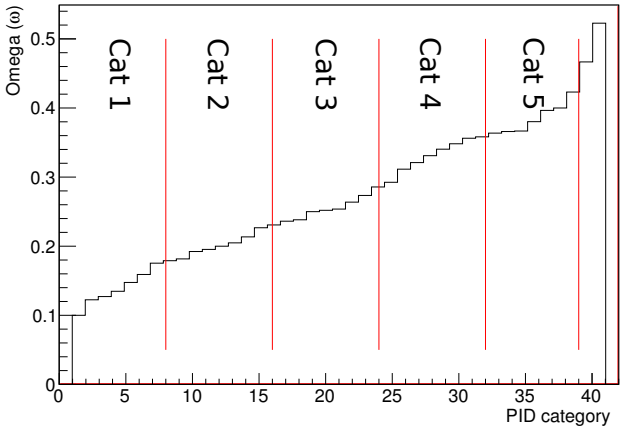


FIGURE 5.11: Wrong tag fraction according to the active tagger. The lines split the sample into 5 tagging categories with the corresponding average mistag. Events with a high ω are considered untagged.

5.5 Event-per-event mistag

In order to best exploit the tagging information and to obtain the best possible results in CP analysis, an event-by-event mistag probability is used to weight the events accordingly (being equivalent to have a category for each event). The effective dilution is then calculated by summing the mistag probabilities on all signal events $\sum_i (1-2\omega_i)^2/N$ [98], with the background events being substracted (for all channels the bakground was substracted using the *sPlots* procedure [93]) and ω calibrated by Eq. 5.3. Table 5.12 reports the event-by-event tagging power

$B_s^0 \to J/\psi\phi$			
	ϵ_{tag} (%)	ω (%)	ϵ_{eff} (%)
cat 1	22.54±0.13	42.4±0.4	0.52±0.05
cat 2	14.47±0.11	36.2±0.4	1.10±0.06
cat 3	8.05 ±0.09	27.9±0.5	1.57±0.08
cat 4	5.23 ±0.07	22.9±0.6	1.53±0.07
cat 5	3.52 ±0.06	15.9±0.6	1.63±0.07
average	53.81±0.2	34.2±0.2	5.02±0.15
combined	53.81±0.2	32.8±0.2	6.35±0.15

TABLE 5.11: Performance in $B_s^0 \to J/\psi\phi$ decays after the L0, obtained combining the taggers using the PID approximation. To be compared with Table 5.1.

obtained using the calibration parameters determined with the $B^+ \to J/\psi K^+$ events, as reported in Table 5.4, for three control channels. The results on the tagging power are compatible among the channels containing a J/ψ meson. The higher value for $B^0 \to D^{*-}\mu^+\nu_\mu$ is mainly due to the higher tagging efficiency originated from the different trigger selection. We underline that the use of the mistag probability allows to calculate the effective efficiency on any set of selected B events, also for non flavour-specific channels. There is no need to correct the mistag estimated in a control channel in case of phase space differences with the CP channel.

Channel	ϵ_{tag} (%)	$\mathcal{D} = (1-2\omega)$ (%)	ω (%)	$\epsilon_{tag}\mathcal{D}^2$ (%)
$B^+ \to J/\psi K^+$	27.3 ± 0.1	27.8 ± 0.6	36.1 ± 0.3	2.10 ± 0.08
$B^0 \to J/\psi K^*$	26.7 ± 0.2	28.0 ± 0.6	36.0 ± 0.3	2.09 ± 0.09
$B^0 \to D^{*-}\mu^+\nu_\mu$	30.5 ± 0.1	29.0 ± 0.6	35.6 ± 0.3	2.53 ± 0.10

TABLE 5.12: Tagging efficiency, event-by-event tagging dilution, event-by-event effective mistag and event-by-event tagging power for $B^+ \to J/\psi K^+$, $B^0 \to J/\psi K^*$ and $B^0 \to D^{*-}\mu^+\nu_\mu$ events, after background is substracted with the *sPlots* technique. 370 pb^{-1} were used.

These results should be compared with the *average* and *combine* lines from Tables 5.8, 5.9 and 5.10. All the results are compared in Table 5.13 confirming that the use of the event-per-event mistag maximizes the tagging performances.

Channel	$\epsilon_{eff}^{average}$ (%)	$\epsilon_{eff}^{combine}$ (%)	$\epsilon_{eff}^{evt-p-evt}$ (%)
$B^+ \to J/\psi K^+$	1.69 ± 0.1	2.07 ± 0.11	2.10 ± 0.08
$B^0 \to J/\psi K^*$	1.24 ± 0.20	1.57 ± 0.22	2.09 ± 0.09
$B^0 \to D^{*-}\mu^+\nu_\mu$	1.58 ± 0.06	2.05 ± 0.06	2.53 ± 0.10

TABLE 5.13: Comparison of the tagging power on $B^+ \to J/\psi K^+$, $B^0 \to J/\psi K^*$ and $B^0 \to D^{*-}\mu^+\nu_\mu$ events with 370 pb^{-1} using the average, the combined and the event-per-event mistag.

5.6 Conclusion

In this chapter we have detailed how the different flavour tagging algorithms are combined in a single decision and mistag estimation. To avoid an underestimation of the combined mistag, a calibration has to be applied to correct possible correlations between the taggers. There are also several other strategies to combine the taggers, that are additionally explained and tested on MC. In the end, when the final combined decision is obtained and the calibration applied, there are several ways to compute the tagging power that can be used: the event per event is the one used by default in the physics analysis, as the ω is already calibrated. But other methods like the average mistag or the use of tagging categories could be used as well.

Chapter 6

Systematics

In order not to compromise the statistical precision obtainable on any CP violation measurement at LHCb, a good control of the systematic uncertanties must be achieved. In this chapter the main possible sources of systematic uncertainty are discussed.

6.1 Systematic uncertainties

CP measurements require the reconstruction of the final state and, most often, the determination of the flavour in the initial state. Time dependent rates, are combined into asymmetries from which CKM phases can be extracted. Certain charge and flavour dependent effects may exist, which can indeed bias the measurement. The most important ones are:

- **Production asymmetries**
  The initial fraction of b and $\bar{b}$ hadrons is not expected to be the same [99]. Furthermore, the production asymmetry can be different for each hadron specie. The fractions of B^0_d, $\overline{B^0_d}$, B^0_s, $\overline{B^0_s}$, B^+ and B^- mesons per event denoted by f_0, $\overline{f_0}$, f_s, $\overline{f_s}$, f_+ and f_- will have an error and it has to be taken into account.

- **Charge dependence**
  All flavour tagging methods rely on measuring the charge of one or more selected tracks. If the track reconstruction efficiency, or particle assignment, has a charge dependence, it will result in a difference in the tagging efficiency for b and $\bar{b}$ hadrons. An asymmetric tagging efficiency can develop from effects such as differences in interaction cross sections for K^+ and K^- with matter. In LHCb we can split the data according to the magnet polarity in order to check whether possible asymmetries of the detector efficiency or of the alignment accuracy can introduce a difference in the tagging performance.

Additionally, due to this different particle/antiparticle interaction with matter or to possible detector asymmetries, the calibration of the mistag probability of different B flavour might be different. In this case a systematic uncertainty has to be considered, unless the difference is explicitly taken into account while fitting the CP asymmetries. These systematic errors can be measured by splitting the sample into two subsamples according to the signal flavour, determined by the reconstructed final state, or the tagging decision.

- **Other systematics**
  There are several other effects that can affect the tagging performances in both, tagging efficiency and mistag probability. Systematic can arise from different running conditions: the number of reconstructed primary vertices, the track multiplicity. Also from a different p_T spectra of the signal B, caused by different trigger and offline selections.

In the following sections only the main sources of systematic effects related to the mistag are reviewed. These possible systematics must be taken into account when the average ω is used. For most of the CP analyses where the per-event mistag is used, systematic uncertainties associated to the calibration parameters should be taken into account.

6.2 Dependence of the OS tagging on event properties

Dependence on the flavour of the signal B and on the tag decision

With 1 fb^{-1}, we studied if there was any flavour dependence in the tagging performances by comparing the values obtained by splitting the samples of different control channels according to the B flavour and the sign of the tagging decision [97]. In Table 6.1 the results obtained are compared separately for the OS combination.

The differences of K^+/K^- interactions in the detector material and kaon production in secondary interactions could have a non negligible effect on tagging, both on the efficiency and on the mistag. Studies performed with 2011 data were not conclusive due to the limited statistical error.

In Table 6.1 are reported in bold the values that show a significative ($> 2.5\sigma$) difference. They correspond to the mistag measured for different tag decisions, in the case of $B^+ \to J/\psi K^+$ and $B^0 \to J/\psi K^*$ channels. It is interesting to notice that the discrepancy in the mistag fraction for $B^+ \to J/\psi K^+$ and $B^- \to J/\psi K^-$ can be explained by an asymmetry in the statistics of the two samples B^+ and B^-. For the $B^+ \to J/\psi K^+$ channel it is possible to compute the value equalizing the number of B^+ and B^- in the initial sample (marked by the label ^{EQ}). A similar check cannot be done with the B^0 channels as they oscillate.

OS		ϵ_{tag}[%]	ω [%]	$\epsilon_{tag}(1-2\omega)^2$ [%]		ϵ_{tag}[%]	ω [%]	$\epsilon_{tag}(1-2\omega)^2$ [%]
channel	$\bar{b}$-meson	flavour of the signal B (at decay)			b-meson	flavour of the signal B (at decay)		
$B^+ \to J/\psi K^+$	$J/\psi K^+$	18.85±0.11	35.0±0.3	1.70±0.08	$J/\psi K^-$	18.85±0.12	34.1±0.3	1.92±0.08
$B^0 \to J/\psi K^*$	$J/\psi K^{*0}$	19.4±0.2	36.4±0.7	1.44±0.17	$J/\psi \bar{K}^{*0}$	19.5±0.2	36.1±0.8	1.5±0.2
$B^0 \to D^-\pi^+$	$D^+\pi^-$	23.20±0.16	33.9±0.6	2.42±0.17	$D^-\pi^+$	23.09±0.17	34.3±0.6	2.27±0.16
channel	$\bar{b}$-meson	tag decision			b-meson	tag decision		
$B^+ \to J/\psi K^+$	+1	9.39±0.06	**33.6±0.3**	1.01±0.05	-1	9.46±0.06	**35.5±0.3**	0.80±0.04
$B^+ \to J/\psi K^{+EQ}$	+1	9.33±0.06	34.4±0.3	0.91±0.04	-1	9.51±0.06	34.7±0.3	0.90±0.04
$B^0 \to J/\psi K^*$	+1	9.71±0.09	**34.7±0.7**	0.91±0.1	-1	9.79±0.09	**37.8±0.7**	0.58±0.08
$B^0 \to D^-\pi^+$	+1	23.20±0.16	33.9±0.6	2.42±0.17	-1	23.09±0.17	34.3±0.6	2.27±0.16
channel	$\bar{b}$-meson	Magnetic field			b-meson	Magnetic field		
$B^+ \to J/\psi K^+$	up	18.79±0.13	34.3±0.4	1.86±0.10	down	18.89±0.11	34.7±0.3	1.77±0.08
$B^0 \to J/\psi K^*$	up	19.3±0.3	36.2±0.8	1.47±0.19	down	19.6±0.2	36.3±0.7	1.47±0.16
$B^0 \to D^-\pi^+$	up	22.99±0.18	34.1±0.6	2.33±0.18	down	23.26±0.15	34.1±0.5	2.36±0.15

TABLE 6.1: Performance of the OS taggers for the control channels analyzed with 1 fb^{-1}. The results correspond to the average values obtained applying the cut η_c <0.42. See Ref. [97].

Dependence of the OS tagging performances on the magnet polarity

The different interaction cross-section between particle and antiparticle in the detector material could have a non negligible effect on tagging, both on efficiency and on mistag. An example could be the differences between K^- and K^+ production in secondary interactions affecting the SSK, the OSK and vertex charge taggers. A possible systematic due to asymmetries in the detector acceptance and/or efficiency can be studied by comparing the opposite polarities of the magnetic field. With the statistics available, no asymmetry is observed comparing the tagging performances with the two B-field polarities (see Table 6.1).

Dependence of the OS tagging performances on the p_T of the signal B

The tagging performances were also evaluated in bins of p_T. Figure 6.1 shows the p_T distribution for different control channels. In Figure 6.2 the tagging performance on $B^+ \to J/\psi K^+$ events are evaluated in four bins of p_T defined by the ranges: 0-3 GeV/c, 3-6 GeV/c, 6-10 GeV/c and $>$ 10 GeV/c. The measured OS mistag only shows a very mild dependence on the p_T of the signal B, while the tagging power increases with p_T due to the increase of the tagging efficiency [97].

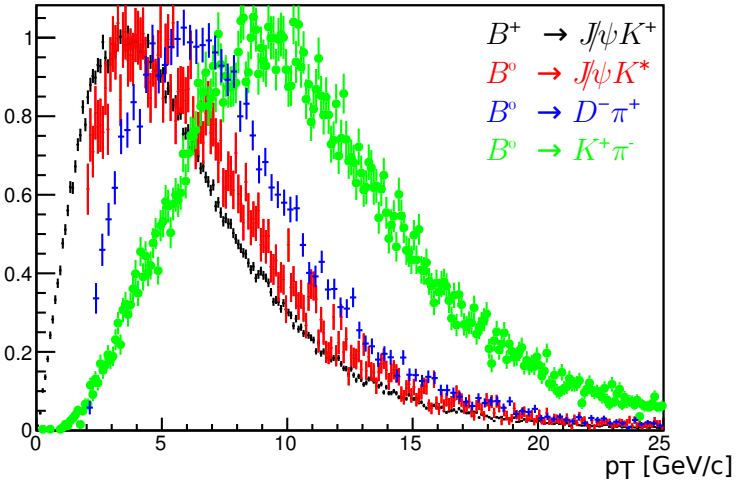


FIGURE 6.1: p_T distribution for $B^+ \to J/\psi K^+$ (black), $B^0 \to J/\psi K^*$ (red), $B^0 \to D^-\pi^+$ (blue) and $B^0 \to K^+\pi^-$ (green) channels.

The invariance of the mistag value on the p_T is an important feature that allows to use the measured mistag in a control channel as reference value for another one without the need of accounting for different p_T distributions. This is not the case for SS taggers, where the different offline and trigger selections can make a correction of the phase space necessary.

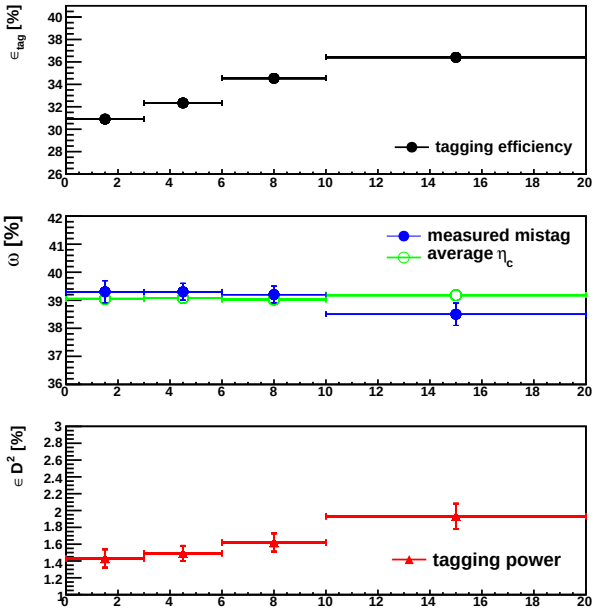


FIGURE 6.2: OS tagging efficiency (top), mistag fraction and average predicted mistag probability (middle), and tagging power (bottom) as a function of the B signal p_T for $B^+ \to J/\psi K^+$ channel. The data sample correspond to the full sample, $0 < \eta_c < 0.5$.

Dependence of the OS tagging performances on the number of reconstructed primary vertices

The performance on reconstruction and particle identification depends on the event multiplicity. As a consequence, also the tagging performance can be affected. The tagging performances for high multiplicity events on data was first studied with 35 pb^{-1} from the 2010 LHCb physics run [86]. The performances calculated for different numbers of primary vertices (nPV) and track multiplicity were compatible within statistical errors. However, the results seemed to indicate a slight deterioration of the tagging power with the increasing number of collisions, as it was also found in MC.

With 1 fb^{-1} a more precise analysis could be performed thanks to the higher statistics. In Table 6.2 we show the performances of the OS tagging, splitting the data samples according to the number of reconstructed primary vertices: nPV=1, 2 or >2. The results confirm the loss of tagging power with the increasing number of interactions because of the combination of a decrease of the tagging efficiency and a slight increase of the mistag fraction. The same effect can be seen when the number of reconstructed "long tracks" increases [97].

nPV	ϵ_{tag}[%]	ω [%]	$\epsilon_{tag}(1-2\omega)^2$ [%]
$B^+ \to J/\psi K^+$channel			
1	20.78±0.14	34.3±0.4	2.05±0.12
2	19.15±0.13	34.7±0.4	1.80±0.10
>2	15.98±0.15	34.7±0.5	1.50±0.11
$B^0 \to J/\psi K^*$channel			
1	21.2±0.2	35.3±0.8	1.83±0.22
2	19.4±0.2	37.0±0.9	1.31±0.20
>2	16.4±0.2	36.7±1.1	1.16±0.21
$B^0 \to D^-\pi^+$channel			
1	25.5±0.2	33.7±0.6	2.73±0.21
2	22.9±0.2	34.0±0.6	2.34±0.19
>2	23.0±0.3	35.1±0.9	2.05±0.24

TABLE 6.2: Average performance (with $\eta_c < 0.42$) of the OS taggers for different number of reconstructed primary vertices for 1 fb^{-1} on the control channels analyzed.

6.3 Systematics on the OS mistag probability

For a large variety of time-dependent asymmetry measurements, involving different B decay channels, a calibrated event-per-event mistag probability can be used. For these analyses, it is important to determine also the systematic uncertainty on the observable so that it can be propagated to the physics result. The systematic uncertainty on the p_0 and p_1 parameters defined in Section 4.4 has been studied in order to be used for other physics analyses.

In order to evaluate the systematic uncertainties, the calibration procedure on $B^+ \to J/\psi K^+$ events has been repeated under different conditions. Several checks were performed:

- The data sample was split according to the magnet polarity, in order to check whether possible asymmetries of the detector efficiency or of the alignment accuracy introduced a difference in the tagging calibration.
- The data sample was split according to the signal flavour, as determined by the reconstructed final state, or depending on the tag decision to take into account for difference on particle/antiparticle interaction with matter or possible detector asymmetries.
- The data sample were split according to the number of reconstructed primary vertices (nPV), in order to check whether the event multiplicity introduced a difference in the tagging calibration.
- The fit model was changed in different ways: fitting only the mass distribution, changing the decay time acceptance function or using only one exponential for the time distribution of the background. Moreover, the distribution of the mistag probability in the fit model $\mathcal{P}(\eta)$ has been varied either by assuming the signal and background distributions to be

equal or by swapping them. In this way, possible uncertainties related to the fit model are considered.

- Additionally, the stability of the calibration parameters has been verified over the run period and in different bins of transverse momentum of the signal B.

With 370 pb^{-1} integrated luminosity a small difference in the calibration parameters was observed when changing the magnetic field polarity and the flavour of the signal B. A systematic was also observed when changing the $\mathcal{P}(\eta)$ distribution. The sum in quadrature of the mentioned contributions was taken as estimate of the systematic uncertainty being $\delta p_0 = \pm 0.009$ and $\delta p_1 = \pm 0.07$. The calibration parameters were exported together with their statistical and systematic errors to other channels to measure B oscillations and to perform the first CP violation measurements at the LHCb experiment, giving in some cases, the best measurement ever. These first physics results are summarized in Appendix B.

With 1 fb^{-1}, in addition to the systematic induced by fitting the $\mathcal{P}(\eta)$ distribution in different ways as mentioned above, a systematic deviation can be observed when the sample is split according to the initial flavour of the signal B. The same effect is visible if the results are split according to the tagging decision. Since this effect is correlated to the previous one, only the largest is considered for the final computation of the systematics to avoid double counting. Moreover, there is an additional deviation on the p_0 parameters between the $B^0 \to J/\psi K^*$ and $B^0 \to D^{*-}\mu^+\nu_\mu$ channels and the $B^+ \to J/\psi K^+$ one. This is the dominant source of the systematic uncertainty. In Table 6.3 the systematic uncertainties on the OS calibration parameters are summarized. The final systematic uncertainty was determined by the sum in quadrature of the different contributions.

Systematic effect	δp_0	δp_1
Fit model assumptions $\mathcal{P}(\eta)$	$< \pm 0.001$	± 0.012
B-flavour	± 0.005	–
Control channel dependence	± 0.0075	–
Total	± 0.009	± 0.012

TABLE 6.3: Systematic uncertainties on the calibration parameters p_0 and p_1 obtained with 1 fb^{-1} of $B^+ \to J/\psi K^+$ events.

Despite the tagging performance depends on the number of reconstructed primary vertices, the calibration is independent. The data sample was also split according to the number of reconstructed primary vertices (nPV) in order to check whether a difference in the event reconstruction introduces a difference as well in the tagging calibration. The calibration parameters were in agreement with the reference one within the statistical uncertainty, and for this reason no systematic uncertainty was assigned to this source. This behaviour is due to the fact that

one of the neural net inputs is in fact the nPV. The distribution of the predicted mistag depends indeed on the number of primary vertices as it is shown in Figure 6.3.

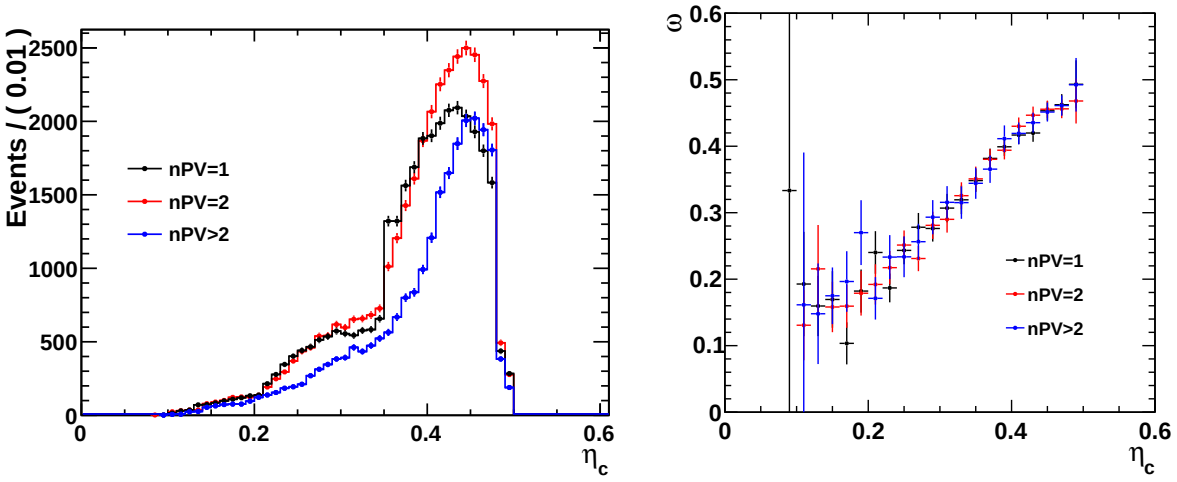


FIGURE 6.3: Distribution of the calibrated mistag probability (left) and calibration plot (right) for different number of primary vertices in $B^+ \to J/\psi K^+$channel.

New results are expected from the physics analyses being performed with 1 fb^{-1} from the whole 2011 physics run. Some of these will need to propagate the uncertainties in the calibration of the mistag probability, for that, the parameters to be used are:

$$\begin{aligned} p_0 &= 0.392 \pm 0.002(\text{stat}) \pm 0.009(\text{syst}) \\ p_1 &= 1.035 \pm 0.021(\text{stat}) \pm 0.012(\text{syst}) \\ <\eta_c> &= 0.391 \qquad \rho(p_0, p_1) = 0.12 \end{aligned} \tag{6.1}$$

6.4 Systematic studies on Same Side Kaon tagging

Different studies were performed to extract the uncertainties in the calibration parameters of the SS kaon with the $B_s^0 \to D_s^- \pi^+$ channel. First, the SSK calibration parameters were split according to the decay flavour of the B_s meson, the magnetic field and the charge of the tagging kaon. No significant differences were found in the two samples and no systematic error was assigned for these effects. Additionally, a systematic error of $\delta p_0 = \pm 0.004$ and $\delta p_1 = \pm 0.02$ was assigned due to differences in the fit model as described in Section 4.7.4, and an uncertainty of $\delta p_0 = \pm 0.006$ and $\delta p_1 = \pm 0.01$ was also assigned due to the scaling factor of the proper time resolution. Taking into account the different uncertainties, the resulting calibration for 1 fb^{-1} is:

$$
\begin{aligned}
p_0 &= 0.350 \pm 0.015(\text{stat}) \pm 0.005(\text{syst}) \\
p_1 &= 0.51 \pm 0.16(\text{stat}) \pm 0.02(\text{syst}) \\
&< \eta_c >= 0.324
\end{aligned}
\tag{6.2}
$$

These systematic uncertainties on the SSK are still under study and will possibly change in the near future. They are described in Ref. [95].

Chapter 7

Summary of performances

The identification of the quark content of the signal B meson at the production time is mandatory to measure B oscillations or for precise CP violation measurements. Thanks to the flavour tagging the LHCb experiment has been able to measure the first B oscillations and obtain different CP violation measurements, some of them with the higher precision. In this chapter we summarize the flavour tagging performances in different control channels for different luminosities.

7.1 Flavour Tagging in control channels

One of the most important measurements to be made at LHCb where flavour tagging plays a very important role is the ϕ_s measurement on the $B_s^0 \to J/\psi\phi$ channel. For that purpose, the flavour tagging algorithms were optimized with J/ψ channels. Although the performances should be similar between channels, small differences in trigger, selection, etc. can appear. Special crosschecks have been made on $B^0 \to D^{*-}\mu^+\nu_\mu$ to validate the tagging and prove its universality. In what follow we will refer to the analyses carried out with different datasets.

Integrated Luminosity of 35 pb^{-1}

The first measurements of the flavour tagging performances are detailed in [86] and were obtained with the first 35 pb^{-1} from the 2010 physics run. At that time, only the average and the combined performance were calculated, due to the lack of statistics the systematic uncertainty was not introduced. In Tables 7.1, 7.2 and 7.3 there is a summary of the tagging performances for the $B^+ \to J/\psi K^+$, $B^0 \to D^{*-}\mu^+\nu_\mu$ and $B^0 \to J/\psi K^*$ control channels measured with the optimized cuts from Tables 4.17, 4.18 and 4.19 labeled as *Moriond 2011*. Additionally in Figure

7.1 you can see the first B^0 oscillation observed with $B^0 \to D^{*-}\mu^+\nu_\mu$ and $B^0 \to J/\psi K^*$ channels.

$B^+ \to J/\psi K^+$	$\epsilon_{tag}(\%)$	ω (%)	$\epsilon_{tag}(1-2\omega)^2$ (%)
OS average (η_c <0.42)	15.4±0.4	33.3±1.2	1.71±0.29
OS combine (η_c <0.44)	15.4±0.3	32.2±1.2	1.97±0.31
OS + SSπ average (η_c <0.42)	22.7±0.4	35.5±1.2	1.92±0.30
OS + SSπ combine (η_c <0.44)	23.0±0.5	33.9±1.2	2.38±0.33

TABLE 7.1: Tagging performance for $B^+ \to J/\psi K^+$ channels. OS and SSπ + OS combination are shown. The uncertainties reflect the available statistics of 2010 data.

$B^0 \to D^{*-}\mu^+\nu_\mu$	$\epsilon_{tag}(\%)$	ω (%)	$\epsilon_{tag}(1-2\omega)^2$ (%)
OS average (η_c <0.44)	18.2±0.2	34.3±0.8	1.79±0.18
OS combine (η_c <0.44)	18.3±0.2	33.6±0.8	1.97±0.18
OS + SSπ average (η_c <0.44)	29.1±0.2	35.8±0.8	2.36±0.26
OS + SSπ combine (η_c <0.44)	28.9±0.2	34.2±0.8	2.87±0.32

TABLE 7.2: Tagging performance for $B^0 \to D^{*-}\mu^+\nu_\mu$ channels. OS and SSπ + OS combination are shown. The uncertainties reflect the available statistics of 2010 data.

$B^0 \to J/\psi K^*$	$\epsilon_{tag}(\%)$	ω (%)	$\epsilon_{tag}(1-2\omega)^2$ (%)
OS average (η_c <0.44)	15.7±0.6	33.1±3.0	1.79±0.71
OS combine (η_c <0.44)	15.8±0.7	30.0±6.6	2.52±0.82
OS + SSπ average (η_c <0.44)	25.9±0.8	37.0±2.4	1.75±0.70
OS + SSπ combine (η_c <0.44)	26.1±0.9	33.6±5.1	2.82±0.87

TABLE 7.3: Tagging performance for $B^0 \to J/\psi K^*$ channels. OS and SSπ + OS combination are shown. The uncertainties reflect the available statistics of 2010 data.

Thanks to the first optimization and calibration of the flavour tagging algorithms the first measurements of B^0 and B_s^0 oscillation together with the first $\sin 2\beta$ measurement could be performed at the LHCb. These analysis are explained in Appendix B.

Integrated Luminosity of 370 pb^{-1}

In summer 2011, with 370 pb^{-1} a new optimization and calibration was obtained for the tagging algorithms. The tagging performances of the single taggers and of the OS combination, measured after the optimization of the single taggers and the calibration of the mistag probability, are shown in Tables 7.4, 7.5 and 7.6 for the $B^+ \to J/\psi K^+$, $B^0 \to J/\psi K^*$ and $B^0 \to D^{*-}\mu^+\nu_\mu$ channels, respectively. The performance of the OS combination is evaluated in different ways as detailed in Chapter 5.

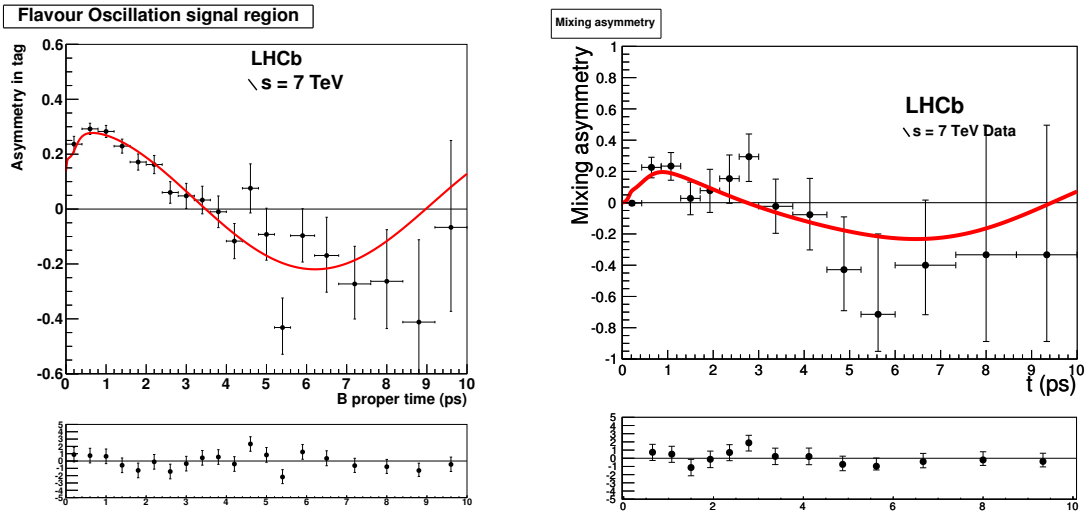


FIGURE 7.1: Mixing asymmetry of $B^0 \to D^{*-}\mu^+\nu_\mu$ (left) and $B^0 \to J/\psi K^*$ (right) events for 35 pb^{-1}, in the signal mass region, as determined by using the combination of all OS and SSπ taggers. The mass difference was fixed at $\Delta m_s = 0.507$ ps^{-1}. The lower plot shows the residuals with respect to the fitted function.

The *average* performance of the OS combination is calculated giving the same weight to each event. Events with a poor predicted mistag probability η_c (larger than 0.42) are rejected. Additionally, to better exploit the tagging information, the tagging performance is combined from independent samples obtained by sorting the data in bins of η_c. The use of such tagging categories, labeled as *combine* in Tables, improves the tagging power by approximately 30% with respect to the average values. The event-per-event value is obtained with the calibration parameters extracted from the $B^+ \to J/\psi K^+$ events, hence the uncertanties reflect the statistical precision of the calibration.

Taggers	ϵ_{tag}[%]	ω [%]	$\epsilon_{tag}(1-2\omega)^2$ [%]
μ	4.78±0.07	29.9±0.7	0.77±0.07
e	2.23±0.05	33.2±1.1	0.25±0.04
K	11.6±0.1	38.3±0.5	0.63±0.06
$Q_{\rm vtx}$	15.1±0.1	40.0±0.4	0.60±0.06
OS average (η_c <0.42)	17.81±0.13	34.6±0.4	1.69±0.10
OS combine	27.3±0.2	36.2±0.5	2.07±0.11
OS event-by-event	-	-	2.10±0.08

TABLE 7.4: Tagging performance in the $B^+ \to J/\psi K^+$ channel for 370 pb^{-1}. Uncertainties are statistical.

The measured tagging performance is similar among the three channels. The differences between the $B^+ \to J/\psi K^+$ and $B^0 \to J/\psi K^*$ results are large in absolute terms, but still compatible

Taggers	ϵ_{tag}[%]	ω [%]	$\epsilon_{tag}(1-2\omega)^2$ [%]
μ	4.84±0.12	34.3±1.9	0.48±0.12
e	2.18±0.08	32.4±2.8	0.27±0.10
K	11.4±0.2	39.6±1.2	0.49±0.13
$Q_{\rm vtx}$	14.9±0.2	41.7±1.1	0.41±0.11
OS average (η_c <0.42)	17.9±0.2	36.8±1.0	1.24±0.20
OS combine	27.1±0.3	38.0±0.9	1.57±0.22
OS event-by-event	-	-	2.09±0.09

TABLE 7.5: Tagging performance in the $B^0 \to J/\psi K^*$ channel for 370 pb^{-1}. Uncertainties are statistical only.

Taggers	ϵ_{tag}[%]	ω [%]	$\epsilon_{tag}(1-2\omega)^2$ [%]
μ	6.08±0.04	33.3±0.4	0.68±0.04
e	2.49±0.02	34.3±0.7	0.25±0.02
K	13.36±0.05	38.3±0.3	0.74±0.04
$Q_{\rm vtx}$	16.53±0.06	41.5±0.3	0.48±0.03
OS average (η_c <0.42)	20.56±0.06	36.1±0.3	1.58±0.06
OS combine	30.48±0.08	37.0±0.3	2.06±0.06
OS event-by-event	-	-	2.53±0.10

TABLE 7.6: Tagging performance in the $B^0 \to D^{*-}\mu^+\nu_\mu$ channel for 370 pb^{-1}. Uncertainties are statistical only.

within three standard deviations, given the large statistical uncertainties of the $B^0 \to J/\psi K^*$ results. Differences between the $B^0 \to D^{*-}\mu^+\nu_\mu$ and $B \to J/\psi X$ results are attributed to the different trigger and selection requirements of the $B^0 \to D^{*-}\mu^+\nu_\mu$ events.

The average B_{sig} p_T is lower in J/ψ channels than in hadronic channels, this is due to the fact that J/ψ channels are more likely to be triggered by the signal muons, where a lower p_T cut is needed with respect to the hadronic channels. Therefore, different p_T distributions are expected for TOS (Trigger On Signal) events between decays including muons and decays including only hadrons. In the case of SS taggers, the mistag has a clear dependence with the B p_T, thus the higher is the p_T, the better is the performance.

Furthermore, the fraction of TIS (Trigger Independent on Signal) events plays a role in the tagging performances. The fraction of TIS events is higher in hadronic decay modes, while J/ψ channels are mostly TOS. The tagging performance is always better in TIS events, where it is more probable to have triggered because of a tagging particle of high p_T. This results in a decrease of ω and an increase of the tagging efficiency.

In case of the SS taggers, the performance of the SSπ was measured on the $B^0 \to D^{*-}\mu^+\nu_\mu$ channel giving $\epsilon_{tag} = 17.08 \pm 0.06\%$, $\omega = 39.3 \pm 0.4\%$, and $\epsilon_{\mathit{eff}} = 0.78 \pm 0.05\%$. For the SS$K$, a subsample of around 580 $B_s^0 \to D_s^-\pi^+$ events were used. The tagging efficiency for this sample was $\epsilon_{tag} = 16.0 \pm 0.6\%$ and the wrong tag fraction $\omega = 33 \pm 5\%$ giving an $\epsilon_{\mathit{eff}} = 1.8 \pm 1\%$.

In Table 7.7 we show the OS event-by-event tagging performances obtained by propagating the statistical and systematic uncertainties of the calibration parameters determined with the $B^+ \to J/\psi K^+$ events. In addition to the values for the three control channels the result obtained for $B_s^0 \to J/\psi\phi$ events is shown. For all channels the background was unfolded using the *sPlot* technique [93]. The results on the tagging power are compatible among the channels containing a J/ψ meson while the higher value for $B^0 \to D^{*-}\mu^+\nu_\mu$ is mainly due to the higher tagging efficiency originating from the different trigger selection.

Channel	$\epsilon_{tag}[\%]$	ω [%]	$\epsilon_{tag}\mathcal{D}^2$ [%]
$B^+ \to J/\psi K^+$	27.3 ± 0.1	$36.1 \pm 0.3 \pm 0.8$	$2.10 \pm 0.08 \pm 0.24$
$B^0 \to J/\psi K^*$	26.7 ± 0.2	$36.0 \pm 0.3 \pm 0.8$	$2.09 \pm 0.09 \pm 0.24$
$B^0 \to D^{*-}\mu^+\nu_\mu$	30.5 ± 0.1	$35.6 \pm 0.3 \pm 0.8$	$2.53 \pm 0.10 \pm 0.27$
$B_s^0 \to J/\psi\phi$	24.9 ± 0.5	$36.1 \pm 0.3 \pm 0.8$	$1.91 \pm 0.08 \pm 0.22$

TABLE 7.7: Tagging efficiency, mistag probability and tagging power for the OS combination calculated from event-by-event probabilities on 0.37 fb^{-1}, for $B^+ \to J/\psi K^+$, $B^0 \to J/\psi K^*$ $B^0 \to D^{*-}\mu^+\nu_\mu$ and $B_s^0 \to J/\psi\phi$ signal OS tagged events. The quoted uncertainties are obtained propagating the statistical (first) and systematic (second) uncertainties on the calibration parameters determined from the $B^+ \to J/\psi K^+$ events.

In Figure 7.2 the fit to the flavour oscillation used to obtain the mistag probability in the $B^0 \to D^{*-}\mu^+\nu_\mu$ and $B^0 \to J/\psi K^*$ channels are shown [96].

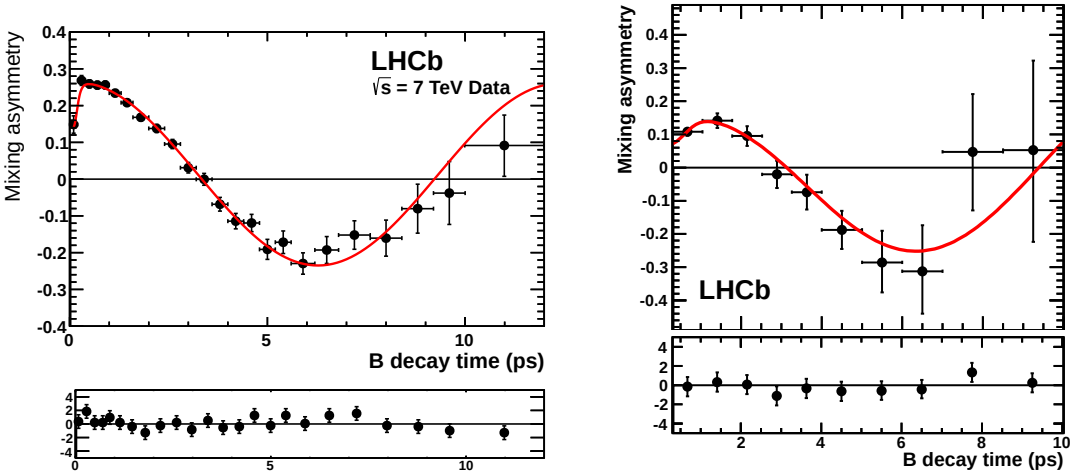


FIGURE 7.2: Mixing asymmetry of $B^0 \to D^{*-}\mu^+\nu_\mu$ (left) and $B^0 \to J/\psi K^*$ (right) events for 370 pb^{-1}, in the signal mass region, as determined by using the combination of all OS taggers. The lower plot shows the residuals with respect to the fitted function.

Integrated Luminosity of 1 fb^{-1}

At the end of 2011, a new optimization and calibration were applied to the whole 1 fb^{-1} collected from the full 2011 physics run. In Tables 7.8, 7.9 and 7.10 the performances are shown for the OS tagged events on the $B^+ \to J/\psi K^+$, $B^0 \to J/\psi K^*$ and $B^0 \to D^-\pi^+$ control channels. The performances were measured with the optimized cuts from Tables 4.17 and 4.19 labeled as *Moriond 2012*. In Figure 7.3 the fit to the flavour asymmetry is shown for $B^0 \to J/\psi K^*$ and $B^0 \to D^-\pi^+$ OS tagged events.

Taggers	$\epsilon_{tag}(\%)$	ω (%)	$\epsilon_{tag}(1-2\omega)^2$ (%)
Single taggers			
μ	5.20±0.04	30.8±0.4	0.77±0.04
e	2.46±0.03	30.9±0.6	0.36±0.07
K	17.67±0.08	39.33±0.24	0.81±0.04
Q_{vtx}	18.46±0.08	40.31±0.24	0.70±0.04
Combination of taggers **Average value**			
OS ($\eta_c < 0.42$)	18.85±0.08	34.5±0.2	1.81±0.05
Combination of taggers **Categories**			
OS cat1	21.72±0.08	43.78±0.22	0.34±0.03
OS cat2	6.30±0.05	35.0±0.4	0.57±0.03
OS cat3	3.43±0.04	26.6±0.5	0.75±0.04
OS cat4	1.30±0.02	20.4±0.7	0.46±0.03
OS cat5	0.40±0.013	15.5±1.2	0.19±0.02
OS sum of categories	33.2±0.11	36.8±0.2	2.31±0.07
Combination of taggers: **event-by-event** $\sum_i (1-2\omega^i)^2/N$ $\omega^i = p_0 + p_1(\eta^i - <\eta_c>)$; no cut on η_c			
OS	33.17±0.09	36.7±0.2	2.35±0.06

TABLE 7.8: Performance of the OS taggers and combination for the $B^+ \to J/\psi K^+$ channel. The results of the tagging power calculated event-by-event is obtained using the final calibration parameters from the $B^+ \to J/\psi K^+$ channel. Uncertainties are statistical only.

For the same side taggers, the set of cuts which optimize the performances are summarized in Tables 4.10 and 4.18. The performances for the SS taggers are summarized in Table 7.11.

Taggers	$\epsilon_{tag}(\%)$	ω (%)	$\epsilon_{tag}(1-2\omega)^2$ (%)
Single taggers			
μ	5.18±0.07	33.8±1.0	0.54±0.07
e	2.49±0.05	36.0±1.4	0.19±0.04
K	18.1±0.1	40.0±0.6	0.72±0.09
$Q_{\rm vtx}$	18.4±0.1	39.9±0.6	0.75±0.09
Combination of taggers **Average value**			
OS ($\eta < 0.42$)	19.0±0.1	36.1±0.5	1.46±0.12
Combination of taggers **Categories**			
OS cat1	21.9±0.13	44.4±0.5	0.27±0.05
OS cat2	6.31±0.07	36.6±0.9	0.45±0.07
OS cat3	3.49±0.06	28.4±1.2	0.65±0.08
OS cat4	1.25±0.03	21.3±1.8	0.41±0.06
OS cat5	0.46±0.02	12.6±2.8	0.26±0.05
OS sum of categories	33.4±0.2	37.6±0.4	2.04±0.14
Combination of taggers: **event-by-event** $\sum_i (1-2\omega^i)^2/N$ $\omega^i = p_0 + p_1(\eta^i - <\eta_c>)$; no cut on η_c			
OS	33.27 ± 0.14	36.7 ± 0.2	2.37 ± 0.06

TABLE 7.9: Performance of the OS taggers and combination for the $B^0 \to J/\psi K^*$ channel. The results of the tagging power calculated event-by-event is obtained using the final calibration parameters from the $B^+ \to J/\psi K^+$ channel. Uncertainties are statistical only.

Taggers	$\epsilon_{tag}(\%)$	ω (%)	$\epsilon_{tag}(1-2\omega)^2$ (%)
Single taggers			
μ	7.47±0.07	30.0±0.7	1.19±0.8
e	3.09±0.05	32.7±1.1	0.37±0.5
K	14.70±0.10	37.9±0.5	0.86±0.7
$Q_{\rm vtx}$	14.23±0.10	37.1±0.5	0.95±0.8
Combination of taggers **Average value**			
OS ($\eta_c < 0.42$)	23.1±0.1	34.1±0.4	2.34±0.12
Combination of taggers **Categories**			
OS cat1	23.62±0.12	43.6±0.4	0.39±0.05
OS cat2	7.67±0.07	35.5±0.7	0.64±0.06
OS cat3	4.59±0.06	39.0±0.9	0.81±0.07
OS cat4	1.98±0.04	22.3±1.2	0.61±0.06
OS cat5	0.74±0.02	13.5±1.8	040±0.04
OS sum of categories	38.60±0.16	38.0±0.3	2.85±0.12
Combination of taggers: **event-by-event** $\sum_i (1-2\omega^i)^2/N$ $\omega^i = p_0 + p_1(\eta^i - <\eta_c>)$; no cut on η_c			
OS	38.51 ± 0.15	38.29 ± 0.04	3.16 ± 0.24

TABLE 7.10: Performance of the OS taggers and combination for the $B^0 \to D^-\pi^+$ channel. The results of the tagging power calculated event-by-event is obtained using the final calibration parameters from the $B^+ \to J/\psi K^+$ channel. Uncertainties are statistical only.

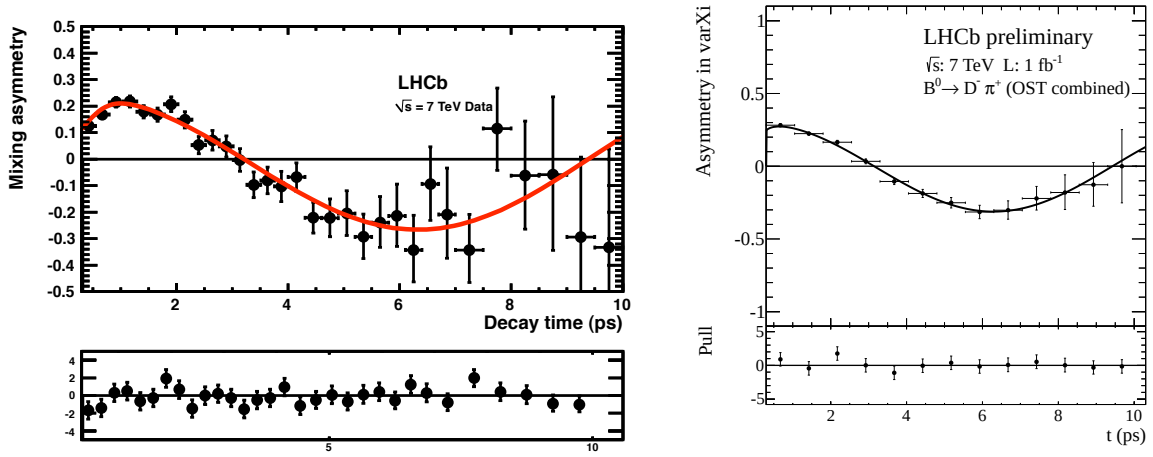


FIGURE 7.3: Mixing asymmetry of $B^0 \to J/\psi K^*$ (left) and $B^0 \to D^-\pi^+$ (right) events for 1 fb^{-1}, in the signal mass region, as determined by using the combination of all OS taggers. The lower plot shows the residuals with respect to the fitted function.

Tagger & Channel	$\epsilon_{tag}[\%]$	ω [%]	$\epsilon_{tag}\mathcal{D}^2$ [%]
SSK $B_s^0 \to D_s^-\pi^+$	16.3 ± 04	35.3 ± 2.1	1.4 ± 0.4
SSπ $B^0 \to J/\psi K^*$	17.6 ± 0.12	39.7 ± 0.5	0.75 ± 0.08
SSπ $B^0 \to D^-\pi^+$	24.08 ± 0.12	39.2 ± 0.4	1.12 ± 0.08

TABLE 7.11: Tagging efficiency, mistag probability and tagging power for the SS taggers from event-by-event probabilities for different control channels. The quoted uncertainties are statistical.

7.2 Physics results

With 35 pb^{-1} the first B^0 - $\bar{B}^0$ oscillations could be measured at LHCb with an oscillation frequency $\Delta m_d = 0.499 \pm 0.032(\text{stat}) \pm 0.003(\text{syst})$ ps^{-1}[36]. Also a first look into CP violation has been done in the time dependent analysis of $B^0 \to J/\psi K^0_s$ by measuring the $\sin 2\beta$. The result is $S_{J\psi K^0_S} = 0.53^{+0.28}_{-0.29}(\text{stat}) \pm 0.08(\text{syst})$ [107]. Both measurements were compatible with the world average. Additionally, the measurement of the B^0_s - $\bar{B}^0_s$ oscillation frequency Δm_s has been made with the best precision ever, giving $\Delta m_s = 17.63 \pm 0.11(\text{stat}) \pm 0.02(\text{syst})$ ps^{-1}[37]. Also the measurement of the CP violating phase ϕ_s in $B^0_s \to J/\psi\phi$ decays has been carried out with 370 pb^{-1} giving the best precision $\phi_s = 0.15 \pm 0.18(\text{stat}) \pm 0.06(\text{syst})$. All these analyses are summarized in the Appendix B.

The new analyses on 1 fb^{-1} with the new optimization and calibration of the tagging algorithms will permit to obtain measurements with an unprecedented precision, and hopefully finding evidences for new physics.

Conclusions

Flavour Physics, and in particular CP violation related effects, plays an important role for the discovery of new physics beyond the Standard Model. For most of the measurements where time-dependent asymmetries are calculated, it is necessary to know the initial flavour of the B meson. This process is known as Flavour Tagging and it is performed at LHCb by means of several algorithms. This document presented a detailed explanation of how the Flavour Tagging is conceived and how it works.

It is based on opposite side (OS) and same side (SS) tagging algorithms to infer the flavour of the signal B meson, either from the identification of the flavour of the opposite B hadron (OS) or by the identification of the charge of the π or K produced in the fragmentation chain of the signal B. The responses of these algorithms are combined later on in order to give a global tagging decision to be used in the CP analyses.

We also developed strategies to optimize and calibrate the tagging algoritms with real data, correcting for the differences between data an MC. They allowed us to maximize the tagging performance and to obtain a reliable mistag estimation used in the CP fits. Both optimization and calibration processes have been successfully applied in many physics analyses with the initial luminosity provided by the LHC.

This led us to important improvements in the flavour tagging algorithms, yielding very good performances in the first year of data. The luminosity collected by the LHCb experiment during 2010 and 2011, allowed us to evaluate the tagging performance in different control channels. The OS calibration parameters obtained from the $B^+ \to J/\psi K^+$ channel can be exported to other channels proving the validity and universality of the tagging response. In the case of SS taggers, the studies are still ongoing, and the final calibration strategy needs to be decided. A number of effects on the tagging performances have been studied like the number of primary vertices, the initial flavour of the signal B and other possible biases due to channel selection and particle interaction with matter, which concur to the evaluation of the systematic error.

So far, thanks to the Flavour Tagging, the LHCb collaboration has been able to perform important measurements and, in some of them, achieving the best precision ever. The B oscillation

frequency Δm_d and Δm_s were measured, the last one with the highest current precision. The $\sin 2\beta$ was measured on $B^0 \to J/\psi K_S$ giving compatible results with the world average. The highest world precision has been achieved in the measurement of $\phi_s^{J/\psi\phi}$, which is one of the CP observables with the smallest theoretical uncertainty, obtaining a compatible result with the Standard Model expectation.

The experience of the 2010 and 2011 data taking has led us to new ideas on how to improve the Flavour Tagging in the LHCb experiment. With the full luminosity and the latest improvements new results of continously increasing precision will be obtained, and hopefully first evidences of new physics beyond the Standard Model may show up.

Appendix A

Control Channels

In this appendix the selections of the events and the specific fits to disentangle signal from background and to extract the wrong tag fraction are explained in detail for each control channel.

A.1 $B^0 \to D^{*-}\mu^+\nu_\mu$ channel

The $B^0 \to D^{*-}\mu^+\nu_\mu$[1] channel is an exclusive semileptonic decay which is selected by requiring that a μ and the decay of $D^{*-} \to \bar{D}^0(\to K^+\pi^-)\pi^-$ originate from a common vertex, displaced with respect to the pp interaction point. The muon and $\bar{D}^0$ transverse momenta are required to have $p_T > 0.8$ GeV/c and $p_T > 1.8$ GeV/c respectively. The selection criteria exploit the long B^0 and $\bar{D}^0$ lifetimes by applying cuts like in the impact parameter of the daughters with respect the reconstructed primary vertex. Additional cuts are applied on the muon and kaon particle identification and on the quality of the tracks and vertices.

In semileptonic decays the four-momentum of the b hadron can not be completely reconstructed due to the undetected neutrino. The lack of a tight B mass and the indetermination on the B momentum are limitations for the selection and also for tagging purposes. To correct for the undetected neutrino, it is usually introduced a so called k-factor, which is the ratio between the momentum of the measured decay products ($D^*\mu$ in this case) and the original B meson momentum, obtained from MC simulations. Then, the B^0 proper-time is given by: $\tau = L_B \times m_{B^0}/pc$, where L_B is the decay length of the reconstructed B^0, m_{B^0} the mass, and p is the corrected B^0 momentum which is defined as $p = p_B^{corr} = p(D^*\mu)/k_{fit}$.

The k-factor has a strong dependence on the invariant mass of the reconstructed decay products. A fit to k in bins of $(D^*\mu)$ invariant mass leads to a B momentum resolution of 16% [100]. To obtain a more reliable k-factor, some other strategies can also be used which are detailed in [88].

[1] Charged conjugate modes are implied.

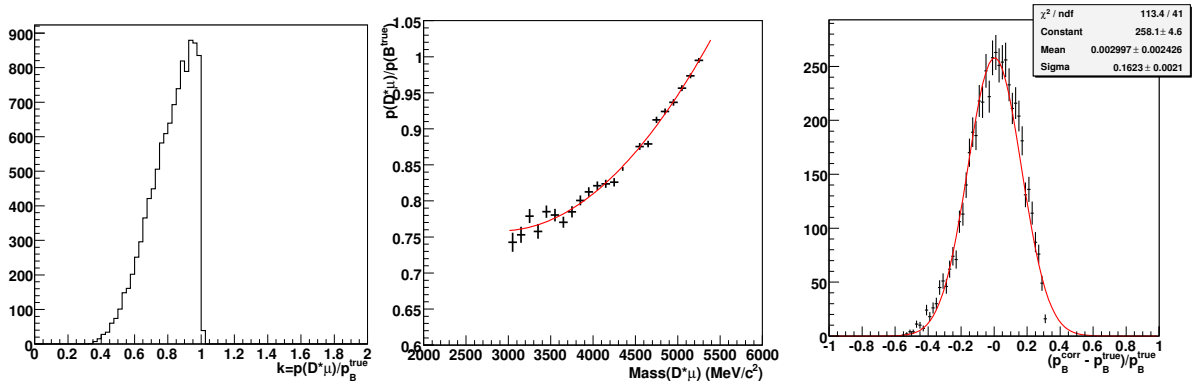


FIGURE A.1: Left: distribution of $k = \frac{p(D^*\mu)}{p_{Btrue}}$ for the selected events. Middle: fit to the k-factor as a function of $D^*\mu$ mass. Right: $(p_B^{corr} - p_B^{true})/p_B^{true}$ where $p_B^{corr} = p(D^{*-}\mu^+)/k_{fit}$ is the corrected B momentum using the k-factor derived from the fit.

Most of them are implemented in the Flavour Tagging to measure the tagging performances in the $B^0 \to D^{*-}\mu^+\nu_\mu$ channel.

To measure the tagging performance on $B^0 \to D^{*-}\mu^+\nu_\mu$ events a complete fit to the flavour oscillation is used. The different components are separated using the information on the $\bar{D}^0$ ($m = m_{K\pi}$) and D^{*-} ($\Delta m = m_{K\pi\pi} - m_{K\pi}$) invariant mass, the B^0 proper-time (t) and the B^0 mixing state (q). An extended unbinned maximum likelihood fit is performed simultaneously on the masses and time distributions of three samples of events: tagged unmixed, tagged mixed and untagged events.

The model describes the selected signal events as $B^0 \to D^{*-}\mu^+\nu_\mu$ decays where the charm mesons are reconstructed in the hadronic modes $D^{*-} \to \bar{D}^0\pi^-$ and $\bar{D}^0 \to K^+\pi^-$. Final states with additional particles which are not explicitly reconstructed are also considered signal. The model also accounts for four type of background events: events from B^+ decaying into the same set of particles plus one or more tracks ($B^+ \to D^{*-}\mu^+\nu_\mu X$), events containing a $\bar{D}^0$ from a b-hadron decay, events containing a D^{*-} which is not originated in a b hadron decay, and background from combinatoric association of tracks.

For the $\bar{D}^0$ and D^{*-} mass peaks, two double Gaussian distributions with a common mean are used, while a specific parametric function from the *RooFit* [103] package is used to describe the Δm distributions of the $\bar{D}^0$ from B and the combinatorial background components. The decay time distribution of the signal consists of mixed, unmixed and untagged events, and it is given by

$$\mathcal{P}^{\rm s}(t,q) \propto \begin{cases} \epsilon_{tag}\; a(t) \left\{ e^{-t/\tau_{B^0}} \left[1 + q(1-2\omega)\cos(\Delta m_d t)\right] \otimes R(t-t') \right\} & \text{if } q = \pm 1 \\ (1-\epsilon_{tag}) a(t) \left\{ e^{-t/\tau_{B^0}} \otimes R(t-t') \right\} & \text{if } q = 0, \end{cases} \tag{A.1}$$

where Δm_d and τ_{B^0} are the B^0–$\bar{B}^0$ mixing frequency and B^0 lifetime. The time acceptance function is denoted by $a(t)$; and $R(t-t')$ is the time resolution model, both extracted from simulation. In total, 28 parameters are free in the fit. A detailed explanation of the fit can be found in [86]. Figure A.2 shows the distribution of the mass and decay time observables used in the maximum likelihood fit for a total sample of $\sim 482\,000$ signal events and a background to signal ratio $B/S \sim 0.14$ in the signal mass region.

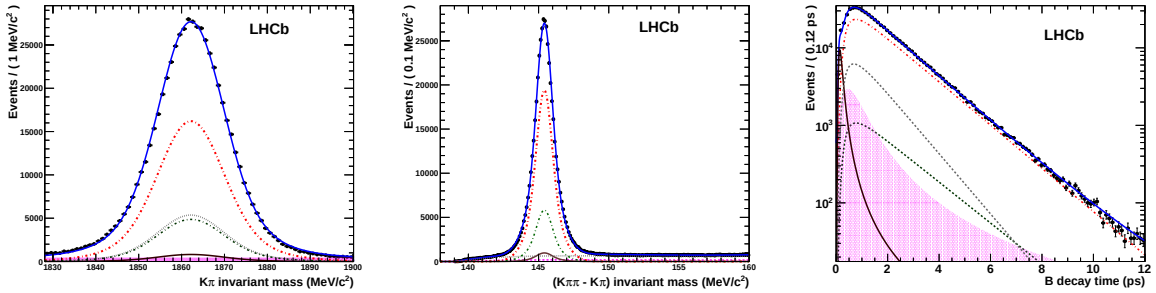


FIGURE A.2: Distributions of invariant mass $K\pi$ (left), mass difference $m(K\pi\pi)-m(K\pi)$ (center) and decay time τ (right) of the $B^0 \to D^{*-}\mu^+\nu_\mu$ events. The black points with errors are data, whereas the blue curve is the fit result. The other lines represent signal (red dot-dashed), $\bar{D}^0$ from B decay background (grey dashed), B^+ background (green short dashed:), and D^* prompt background (magenta solid). Magenta filled area: combinatorial background.

The fit to the flavour asymmetry is needed to obtain the mistag estimation for tagging studies. To do that, the Δm_d was fixed to the reference value $\Delta m_d = 0.507$ ps^{-1} [6]. In Figure A.3 the flavour asymmetry is shown, restricting the plot to the events in the signal mass region. A similar fit without fixing Δm_d was performed in [88] giving compatible results.

A.2 $B^+ \to J/\psi K^+$ channel

The $B^+ \to J/\psi K^+$ candidates were selected by combining $J/\psi \to \mu^+\mu^-$ and K^+ candidates according to the criteria discussed in a dedicated note [104]. The J/ψ mesons are selected by combining two muons with transverse momenta $p_T > 0.5$ GeV/c which form a common vertex of good quality and have an invariant mass in the range $[3030, 3150]$ MeV/c^2. The K^+ candidates are required to have transverse momenta $p_T > 1$ GeV/c, a momenta $p > 10$ GeV/c and to form a common vertex of good quality with the J/ψ candidate with a resulting invariant mass in a window ± 90 MeV/c^2 around the B^+ mass. Additional requirements on the particle identification of muons and kaons are applied to suppress the background contamination. The decay time t and the invariant mass m of the B^+ are extracted from a vertex fit that includes a constraint to the associated primary vertex, and a constraint to the J/ψ mass for the evaluation of the $J/\psi K^+$ invariant mass. In case of multiple candidates per event only the one with the smallest vertex fit χ^2 is considered. Trigger selections with and without requirements on the

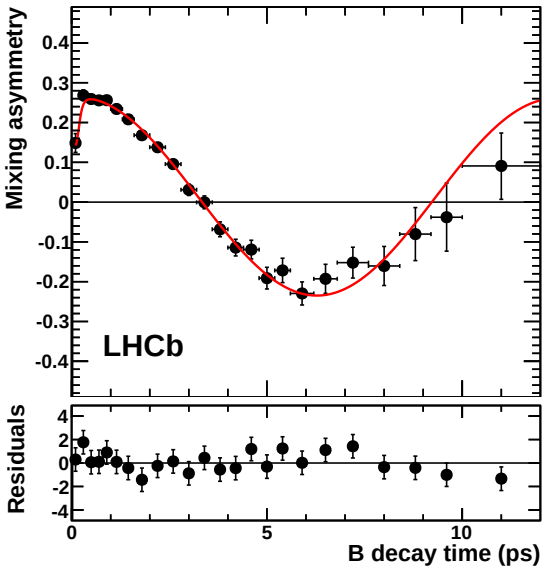


FIGURE A.3: Mixing asymmetry (top) and pull (bottom) as a function of proper-time in $B^0 \to D^{*-}\mu^+\nu_\mu$ events, in the signal mass region, when using the combination of all OS taggers. The B^0 mass difference has been fixed at $\Delta m_d = 0.507 ps$.

IP of the tracks are considered, they are labeled "lifetime biased" (mainly *Hlt1TrackAll* and *Hlt1TrackMuon*) and "lifetime unbiased" (mainly *Hlt1SingleMuonNoIP*, *Hlt1DiMuonNoIP* and *Hlt2UnbiasedJPsi*) respectively.

Having applied this selection, two main sources of background are still present: *prompt* (Pr), given by combination of a K^+ and a J/ψ produced promptly in the pp collision, and *Long lived* (LL), given by the combination of K^+ or a J/ψ produced by a decay of a long living hadron, as partially reconstructed b hadron decays ($H_b \to J/\psi X$), or produced promptly from a different pp collision. Some rejection power is given by the B^+ proper-time: the signal distribution is an exponential with a known lifetime of ∼1.638 ps, while the background is also exponential but has a larger slope. Given that the resolution of the prompt peak is ∼ 0.043 ps, selecting the events with $t > 0.30$ ps is enough to suppress the prompt background component almost completely. The remaining background can be separated from the signal exploiting the different mass distributions, flat for background and peaked on the B^+ mass for signal events. For convenience, both tagged and untagged events were fit simultaneously.

The observables are the B^+ candidate mass m and the tagging decision correctness q_i of each considered tagger i. It takes values: $q_i = +1$ for a right tag, $q_i = -1$ for a wrong tag, and $q_i = 0$ for untagged, respectively. Moreover, for the purposes of the calibration, the probability of mistag η_i, based on the neural net information, can also be used as an observable. In addition, the B^+ candidate decay time t can be also fitted.

The parametrization of all the PDFs is based on the mass, the proper-time and the tagging components. The mass distribution of signal events is described by a double Gaussian (DG) function:

$$\mathcal{S}(m) = f_m \cdot G(m; M_B, \sigma_{m,1}) + (1 - f_m) \cdot G(m; M_B, \sigma_{m,2}) \tag{A.2}$$

while for background events we use a decreasing exponential[2]. The time distribution of the long-lived background component is parametrized with a single exponential, while for signal events it is described by an exponential convoluted with a single Gaussian resolution in the form:

$$\mathcal{S}(t) = \int \exp(-t'/\tau_{B^+}) \cdot G(t - t'; \mu_t^s, \sigma_t^s) \cdot dt' \tag{A.3}$$

For determining the mean tagging parameters of each tagger i, the tagging part of the signal PDF is written as:

$$\mathcal{S}_i^{\rm tag}(q_i) = \begin{cases} \epsilon_i^s \cdot \{1 - \omega_i^s\} & \text{if } q_i = 1, \text{ right tag,} \\ \epsilon_i^s \cdot \omega_i^s & \text{if } q_i = -1, \text{ wrong tag,} \\ 1 - \epsilon_i^s & \text{if } q_i = 0, \text{ untagged,} \end{cases} \tag{A.4}$$

where ω_i^s and ϵ_i^s are the signal mistag fraction and tagging efficiencies. For the calibration of the mistag, the fit parametrization takes into account the probability density function of η in order to extract the p_0 and p_1 calibration parameters. The η distribution is extracted from data for signal and background separately, using events in different mass regions. In that case Eq. A.4 becomes

$$\mathcal{S}_i^{\rm tag}(q_i, \eta_i) = \begin{cases} \epsilon_i^s \cdot \{1 - \omega(\eta)_i^s\} \cdot \mathcal{P}_i^s(\eta) & \text{if } q_i = 1, \text{ right tag,} \\ \epsilon_i^s \cdot \omega(\eta)_i^s \cdot \mathcal{P}_i^s(\eta) & \text{if } q_i = -1, \text{ wrong tag,} \\ 1 - \epsilon_i^s & \text{if } q_i = 0, \text{ untagged,} \end{cases} \tag{A.5}$$

where $\omega(\eta)_i^s = p_0^i + p_1^i \cdot (\eta_i - < \eta_i >)$ and $\mathcal{P}_i^s(\eta)$ is the η distributon. The measured mistag fraction of the background is assumed to be independent on the calculated mistag probability, as confirmed by the data. A more detailed description of the fit can be found in [86]. With 1 fb^{-1} of data 251k signal events were selected with a background to signal ratio $B/S = 0.034$, calculated in a 40 MeV/c^2 mass window around the B^+ mass. Figure A.4 shows the distributions of the reconstructed mass and decay time of the B^+ candidates for OS tagged events, together with the superimposed fit.

[2]The partially reconstructed $B \to J/\psi X$ decays, which would give origin to broadened and shifted mass peaks, are cut away selecting the mass interval $|m - M_{B^+}| < 90$ MeV/c^2

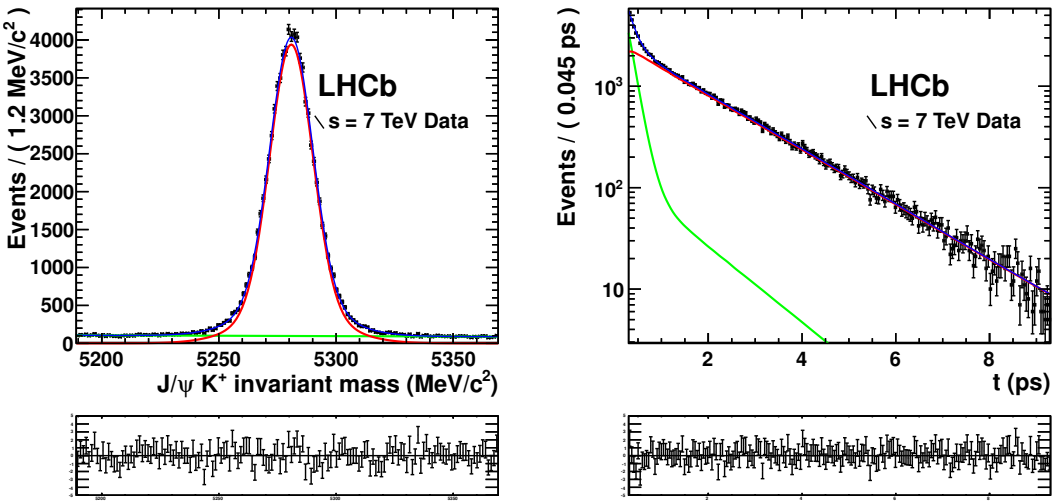


FIGURE A.4: Distributions of the reconstructed B^+ mass (left) and decay time (right) for the OS tagged $B^+ \to J/\psi K^+$ events. Black points are data. The solid blue line, the red line and the green area are: the fit, the signal and the background components, respectively. Lower plots represents the pull distributions.

A.3 $B^0 \to D^-\pi^+$ channel

The selection of the $B^0 \to D^-\pi^+$candidates is described in the document [90]. In particular, only events with a reconstructed decay time $t > 0.2$ ps are analyzed.

With 1 fb^{-1}, the sample contains 105k signal events, and the background to signal ratio is $B/S \approx 0.04$. The most important aspect of the selection is that the $B \to HH$ channels are selected by the same criteria, without requirements on the identification of the final hadrons, which are supposed to be pions. The disentangling of the different specific $B \to HH$ channels is done at analysis level. The analysis of the selected events is described in [90].

A fit to the reconstructed B mass and decay time obtained assuming that the final hadrons are pions is performed. Besides the signal, the fit model accounts for several background contributions due to $B \to HH$ (2-body) channels. Moreover there are also considered additional background contributions due to partially reconstructed B decays to 3-body final states and combinatorial HH background.

The signal mass shape is described by one Crystal-Ball function multiplied by two gaussians. All have the same mean value and different widths. The 2-body and 3-body background contributions are identified on Monte Carlo and mainly due to decay modes in which uncharged particles like π^0 have not been taken into account. The mass fit is performed to ensure that the selected components of hadronic background describe the data in a wide mass range 4950 GeV/c^2 $< m <$ 5750 GeV/c^2. The fit to the mistag values is then performed in a reduced mass

range 5200 GeV/c^2 $< m <$ 5750 GeV/c^2, in order to reduce the background's complexity. In this range, only the signal, a peak from Λ_b decays that is selected due to p/π misidentification and the longlived combinatorial background have to be fit.

A fit to the mass distribution allow us to extract the shape parameters of the mass PDFs which are then fixed in the fit of the mistag probabilities. In Figure A.5 the mass and decay time distributions are shown.

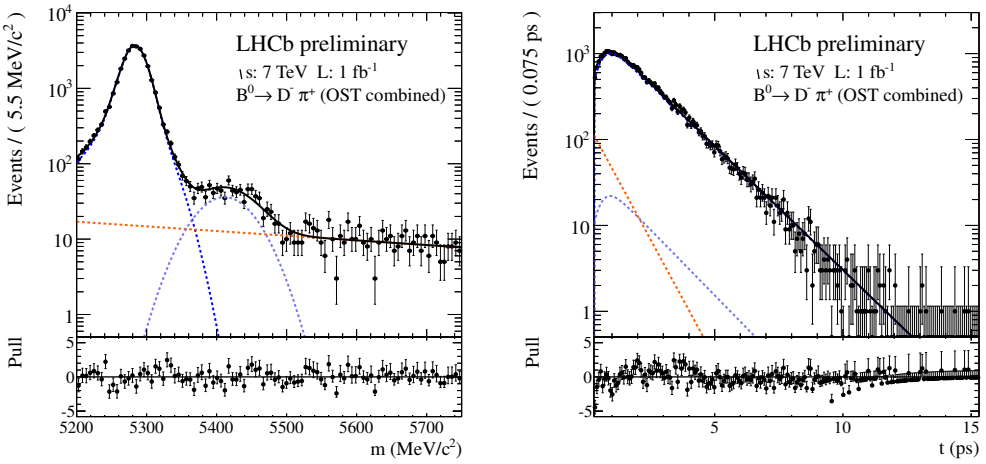


FIGURE A.5: Mass (left) and decay time (right) distributions $B^0 \to D^-\pi^+$events. Black points are data, the solid black line is the fit result. The signal component is represented by the blue line, the different background components are: Λ_b (violet) and combinatorial (orange line).

The decay time acceptance, due to the trigger, stripping and offline cuts is determined using data. The parameters are extracted from a fit to the proper lifetime distribution of the sample as shown in Figure A.5 (right). From a study of the s-weighted datasets it becomes clear that no acceptance has to be aplied to the longlived background. As the statistics for the Λ_b background is very low, the acceptance is compatible with the signal's one. The acceptance function is given by

$$A(t) = \Theta(t - t_0)\frac{at^2 + bt^3 + ct^4}{1 + at^2 + bt^3 + ct^4}, \tag{A.6}$$

where $\Theta(x)$ is the step function and a, b and t_0 are free parameters.

The signal and background decay time distributions are convoluted with the same resolution function, described by a Gaussian of equivalent $< \sigma_t >$=50 fs, extracted from data. The resolution has an insignificant impact on the measurement of the mistag probability.

The mistag fraction is extracted from a fit to all data, with the value Δm_d fixed to the world average [6]. Fig. A.6 shows the time-dependent raw mixing asymmetry in the signal mass region, obtained using the information of the OS tag decision.

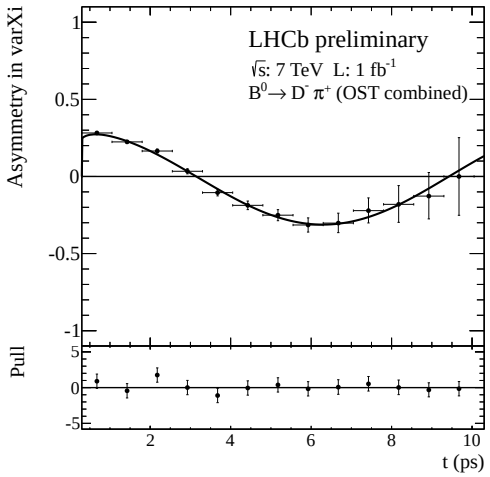


FIGURE A.6: Mixing asymmetry of the $B^0 \to D^+\pi^-$ events in the signal mass region, for all OS tagged events. Black points are data, red solid curve is the asymmetry of the fit function. The lower plot shows the pulls of the residuals with respect to the fitted function.

A.4 $B^0 \to J/\psi K^*$ channel

The $B^0 \to J/\psi K^*$ channel is used to extract the mistag rate through a fit of the flavour oscillation of the B^0 mesons as a function of the decay time. The flavour of the B^0 meson at production time is determined from the tagging algorithms, while the flavour at the decay time is determined from the K^{*0} flavour, which is in turn defined by the kaon charge.

The $B^0 \to J/\psi K^*$ candidates are selected from $J/\psi \to \mu\mu$ and $K^{*0} \to K^+\pi^-$ decays, as it is described in Ref. [104, 105]. To remove the large combinatorial background due to prompt J/ψ production, only events with a reconstructed decay time $t > 0.3$ ps are analysed.

The analysis of the selected events is similar to the $B^+ \to J/\psi K^+$ one, and it is detailed in [97]. The sample of 1 fb^{-1} contains 107k signal events in the range $[5246.5, 5312.5]$ MeV/c^2. The background contribution, with a background to signal ratio $B/S = 0.40 \pm 0.01$ in the same range, is due to partially reconstructed b hadron decays, where a dependence on the decay time is expected (labelled "long-lived" background). We distinguish two long-lived components; the first one corresponds to events where one or more of the four tracks originate from a long-lived particle decay, but where the flavour of the reconstructed K^{*0} is not correlated with a true b-hadron. In the second long-lived background component, one of the tracks used to build the K^{*0} originated from the primary vertex. Both components can be modeled by a decreasing exponential, including an oscillation term in the second case, where the correlation between the K^{*0} flavour and the b is partially lost.

The signal and background decay time distributions are convoluted with the same resolution function, described by a triple Gaussian of equivalent $< \sigma_t >= 50$ fs, extracted from data.

The mass distributions, shown in Figure A.7, are described by a double Gaussian distribution peaking at the B^0 mass for the signal component, and by an exponential with the same exponent for both long-lived backgrounds.

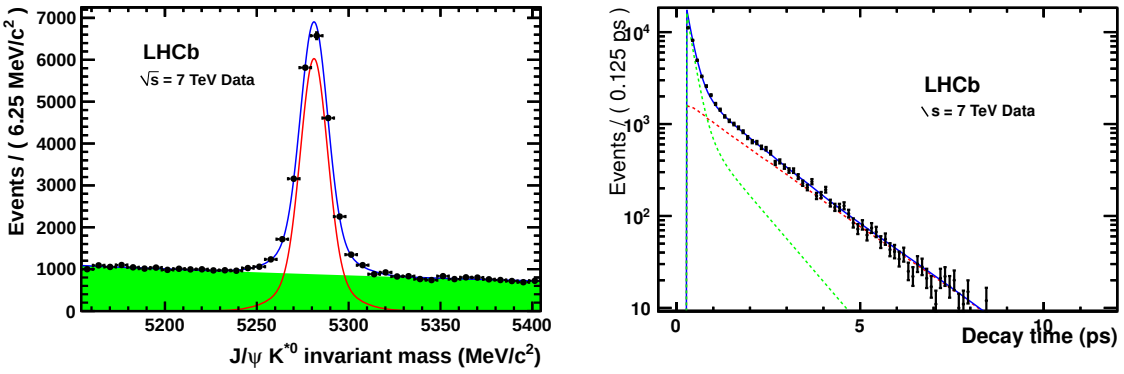


FIGURE A.7: Distribution of the reconstructed mass B^0 (left) and decay time (right) for the $B^0 \to J/\psi K^*$ channel for the events tagged by the OS. Black points are data, the blue curve is the fit result, red curve is the signal component, while the green line represents the sum of the two background components.

The OS mistag fraction is extracted from a fit to all the tagged data, with the values for the B^0 lifetime and Δm_d fixed to the world average [6]. Figure A.8 shows the time-dependent mixing asymmetry in the signal mass region, obtained using the information of the OS tag decision. Letting the Δm_d parameter vary in the fit gives consistent results [90].

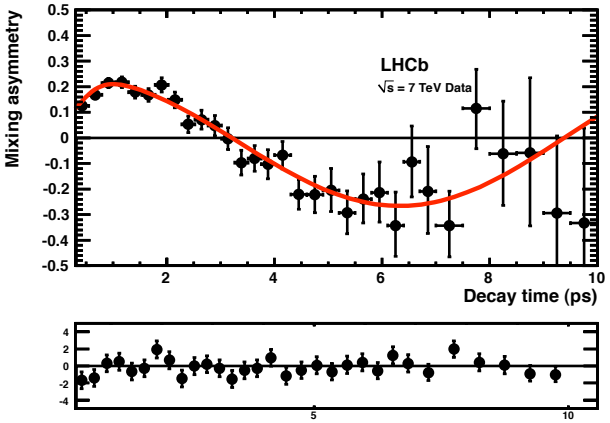


FIGURE A.8: Mixing asymmetry of the $B^0 \to J/\psi K^*$events in the signal mass region, for all OS tagged events. Black points are data, red solid curve is the asymmetry of the fit function. The lower plot shows the pulls of the residuals with respect to the fitted function.

Appendix B

Physics Results

Thanks to the flavour tagging capability, the LHCb experiment has been able to measure the first B oscillations and to obtain different CP violation measurements, some of them with the best precision ever. In this chapter I summarize the first and most important measurements performed with the data from the 2010 and 2011 LHCb physics runs in the flavour and CP fields where the flavour tagging has been of vital importance.

B.1 Physics results using flavour tagging

The good performance of the flavour tagging algorithms has been measured in different control channels obtaining similar and compatible results. These crosschecks prove that the OS calibration parameters obtained from the $B^+ \to J/\psi K^+$ channel can be exported to other channels showing the validity and universality of the tagging algorithms.

Thanks to the flavour tagging the LHCb collaboration has been able to perform important measurements since the first days with only few data. In the following section you will find a summary of the first and most important physics results achieved by the LHCb experiment where the flavour tagging was mandatory. First of all, the tagging was used to measure B^0 oscillations in the $B^0 \to D^-\pi^+$ channel, which allowed the LHCb collaboration to make the first official Δm_d measurement (Section B.1.1). Secondly, the measurement of $\sin 2\beta$ was performed on $B^0 \to J/\psi K_S$ with 37 pb^{-1} of data (Section B.1.2). Already with half of the 2011 data, the Δm_s could be measured in $B_s^0 \to D_s^-\pi^+$ channels obtaining the best measurement ever (Section B.1.3), as well as the ϕ_s phase which was measured with the $B_s^0 \to J/\psi\phi$ channel using only the OS taggers (Section B.1.4).

New results are expected on 2012 with the full data sample collected during the 2011 physics run. The new analyses with 1 fb^{-1} will reach an outstanding precision. The introduction of the

SS kaon tagger, can play an important role, as it can increase the tagging power by nearly a factor 2.

B.1.1 B^0 - $\bar{B}^0$ oscillations

As explained in Chapter 1, weak interactions permit a phenomenon called B oscillations, which allows a B^0 meson to become a $\bar{B}^0$ and viceversa. The B^0 - $\bar{B}^0$ mixing frequency Δm_d has been precisely measured at LEP, at the B factories and at the Tevatron. The world average value from PDG [6] before any LHCb measurement was $\Delta m_d = 0.507 \pm 0.005$ ps^{-1}.

The first measurement of Δm_d at LHCb was performed on 36 pb^{-1} of data taken in 2010. Although it was not competitive with the world best measurements, its main purpose was to serve as a cross check for the flavour tagging, in particular to re-calibrate the flavour tagging algorithms for the Δm_s analysis.

Both opposite and same side tagging algorithms were used in this analysis detailed in [36] where the $B^0 \to D^-\pi^+$ channel was used. In the analysis, the $B^0 \to D^-(K^+\pi^-\pi^-)\pi^+$ signal candidates were selected by a different trigger and passed different stripping and selection cuts than the ones used for the tagging calibration. Re-calibrating on $B^0 \to D^-(K^+\pi^-\pi^-)\pi^+$ decays the flavour tagging calibration obtained in the $B \to J/\psi K^+$ analysis provided a valuable cross-check of the universality of the tagging calibration. Furthermore, the calibration of the opposite side tagging algorithms obtained from this analysis was used as an input for the measurement of Δm_s in the decay $B_s^0 \to D_s^-(K^+K^-\pi^-)(3)\pi$ [37] explained in Section B.1.3.

The selection of the B^0 candidates exploited the long B^0 lifetime, i.e. cuts on the impact parameter of the daughter tracks, the pointing constraint of the B^0, the B^0 lifetime, etc. Additional cuts were applied on particle identification, track and vertex quality, momenta and transverse momenta. Furthermore, the D^- and the B^0 mass have to be consistent with the PDG [6] values within a certain range. About 6000 $B^0 \to D^-(K^+\pi^-\pi^-)\pi^+$ signal candidates were triggered and selected for the blind analysis of Δm_d.

For the analysis it was used an unbinned maximum likelihood where the Probability Density Functions (PDF) for signal and background events factorizes as

$$\mathcal{P} = \mathcal{P}_m(m) \times \mathcal{P}_t^{\epsilon}(t, q|\eta_c) \times \mathcal{P}_{\eta_c}(\eta_c). \tag{B.1}$$

where m refers to the invariant mass of the B^0 candidate, t to its decay time, q with [0,± 1] is the tagging decision and $0 < \eta_c < 1$ is the calibrated mistag probability. $\mathcal{P}_m(m)$ describes the mass distribution. $\mathcal{P}_t^{\epsilon}(t, q|\eta_c)$ describes the distribution of proper decay time according to the B lifetime. It also describes the detector resolution and the measured mixing state. It is

a conditional probability for a given value of the predicted mistag probability η_c. The index ϵ indicates that a decay time acceptance function $\epsilon(t)$ is used to correct for selection biases. $\mathcal{P}_{\eta_c}(\eta_c)$ is the probability distribution for the predicted mistag distribution.

A single Gaussian is used to describe the signal mass distribution; an exponential distribution is used to describe the combinatorial background; and physics backgrounds are derived from Monte Carlo. The fit to the mass distribution was performed in the range [4.80,5.85] GeV/c^2 (Figure B.1 left) determining 5999 $\pm$ 82 $B^0 \to D^-(K^+\pi^-\pi^-)\pi^+$ signal candidates.

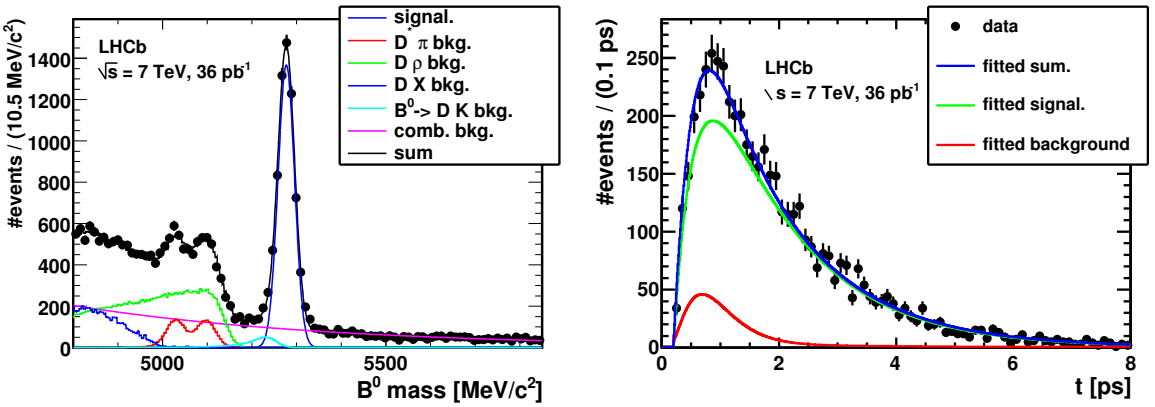


FIGURE B.1: Left: Mass fit in the range [4.80,5.85] GeV/c^2. Only events in the restricted range [5.22,5.85] GeV/c^2 are later used in the lifetime and mixing fit. Right: Result of the lifetime fit.

A double Gaussian (DG) was used to describe the decay time resolution. As the analysed events were collected using a trigger and offline event selection requiring several displaced tracks, an acceptance function $\epsilon(t)$ was introduced to modify the shape of the proper time distribution at low proper times correcting for this effect. The shape of the proper time distribution of the combinatorial background was derived from the high mass sidebands. The lifetime fit can be seen in Figure B.1 (right).

From the fit, the effective tagging efficiency $\epsilon_{eff}^{OS} = \epsilon(1-2\omega)^2$ is found to be 3.4 $\pm$ 0.9 % and the results for the tagging calibration parameters are within the statistical uncertainties in agreement with the expectations from the η_c calibration. The combined tagging performance is $\epsilon_{eff}^{SS+OS} = 4.3 \pm 1.0\%$, reducing the statistical uncertainties on Δm_d from 0.038 to 0.032 ps^{-1} when adding the same side tagger to the analysis.

Different sources of systematic uncertainties were considered: the variation of the decay time resolution in the range 40-63 fs and of the decay time acceptance according to different trigger scenarios, the fixed parameters from the mass and lifetime fit were released and allowed to float in the mixing fit, variations in the PDFs describing the probability distribution of η_c were also considered, a double Gaussian instead of a single Gaussian was used to describe the signal mass PDF and finally, the uncertainty in the z and momentum scale of the experiment were

propagated to the mixing frequency. The total systematic uncertanties were 0.004 and 0.003 ps^{-1} when using the (OS) or the (OS + SS) taggers.

The preliminary value of Δm_d measured using combined opposite and same side taggers on a data set of 36 pb^{-1} collected in 2010 was:

$$\Delta m_d = 0.499 \pm 0.032(\text{stat}) \pm 0.003(\text{syst})\text{ps}^{-1} \tag{B.2}$$

The observed proper time distributions for the different event categories "tagged as mixed", "tagged as unmixed" and "not tagged" are displayed in Figure B.2. The time dependent signal asymmetry (Fig B.2) is defined in the following way:

$$A_{mix}(t) = \frac{N(tagged\ as\ unmixed)(t) - N(tagged\ as\ mixed)(t)}{N(tagged\ as\ unmixed)(t) + N(tagged\ as\ mixed)(t)}, \tag{B.3}$$

where $N(tagged\ as\ (un)mixed)(t)$ refer to the number of B^0 signal candidates which decay at a certain proper time t and are tagged as mixed or unmixed respectively.

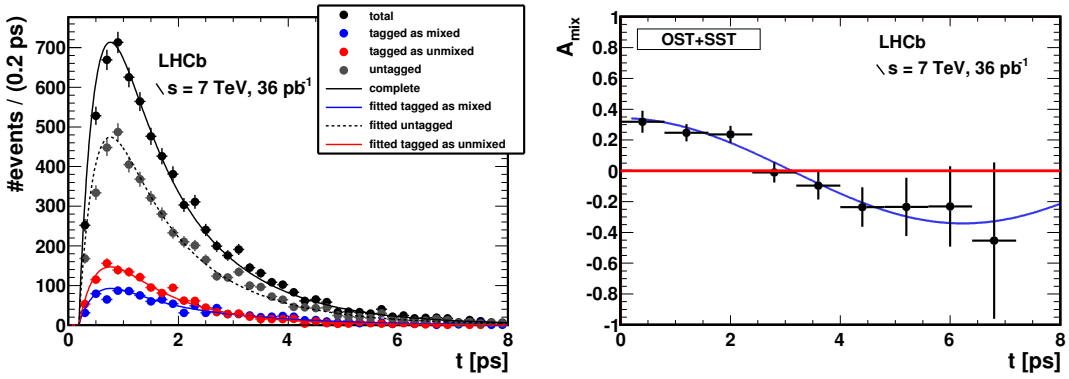


FIGURE B.2: Left: Proper time distribution for different event categories: "tagged as mixed", "tagged as unmixed" and "not tagged". Right: Signal asymmetry as function of proper time. Combined opposite and same side tagging algorithms are used for both plots.

B.1.2 CP violation in the time dependent analysis of $B^0 \to J/\psi K^0_s$

The decay $B^0 \to J/\psi K_S$ is well known as the gold-plated mode for the study of CP violation in the B^0_d meson system. Here, the B^0_d meson decays to a CP eigenstate common to both B^0 and $\bar{B}^0$, the CP violation can appear in the mixing or in the decay, but also in the interference between mixing and decay.

The CKM angle $\sin 2\beta$ is connected to the parameters $S_{J\psi K^0_S}$ and $C_{J\psi K^0_S}$, which govern the time-dependent decay rate asymmetry of the $B^0 \to J/\psi K_S$ decays,

$$\mathcal{A}_{J/\psi K_s} \equiv \frac{\Gamma(\bar{B}^0 \to J/\psi K_s) - \Gamma(B^0 \to J/\psi K_s)}{\Gamma(\bar{B}^0 \to J/\psi K_s) + \Gamma(B^0 \to J/\psi K_s)} = S_{J\psi K^0_S} \sin(\Delta m_d t) - C_{J\psi K^0_S} \cos(\Delta m_d t). \quad \text{(B.4)}$$

Assuming no direct CP violation (the Standard Model predicts negligible direct CP violation for $b \to c\bar{c}s$ decays) $S_{J\psi K^0_S} = \sin 2\beta$ and $C_{J\psi K^0_S} = 0$. Therefore, the measurement of the decay $B^0 \to J/\psi K_S$ has a good sensitivity to the CKM angle $\sin 2\beta$. During the last decade the B Factories Babar and Belle reached an outstanding precision in the measurement of $S_{J\psi K^0_S}$. The world average of $\sin 2\beta$ before the first LHCb measurement had a 3% uncertainty: $\sin 2\beta = 0.673 \pm 0.023$ [106].

The first measurement of $\sin 2\beta$ at LHCb using the $B^0 \to J/\psi K_S$ channel was presented in *Beauty2011* [107] and is detailed in [108].

This first analysis of the $B^0 \to J/\psi K_S$ channel used a data sample corresponding to an integrated luminosity of 35pb^{-1} at a center-of-mass energy of $\sqrt{s} = 7$ TeV collected during the 2010 run. A common selection was used for all the b $\to J/\psi$X decays at the LHCb [109], plus some specific cuts [42]. To increase statistics both the lifetime unbiased and biased ($\sim 20\%$) trigger lines were used.

Only the tagged events are sensitive to $\sin 2\beta$, so a good performance of the tagging is crucial. The OS taggers (electron, muon, kaon and vertex charge) and SS π tagger were used in this analysis where only ~ 280 tagged $B^0 \to J/\psi K_S$ events were selected. The flavour asymmetry that is accessible in $B^0 \to J/\psi K_S$ decays directly depends on the dilution D due to the mistag probability, $D = 1 - 2\omega$, so it is important to have the mistag well established to use as observable the per event mistag probability. The predicted mistag probability was validated on the $B^+ \to J/\psi K^+$ and $B^0 \to J/\psi K^*$ channels as they are self tagging control channels kinematically similar to the signal decay.

The CP violation parameters $S_{J\psi K^0_S}$ and $C_{J\psi K^0_S}$ were extracted through a simultaneous multi-dimensional unbinned extended maximum likelihood fit [42]. The fit was simultaneous in four subsamples: triggered by the lifetime biased/unbiased lines; and by whether or not a tagging decision was available (tagged/untagged). Four observables were also considered: the reconstructed mass m of the B^0_d candidate, its decay time t, the flavour tag decision d, and the calibrated $(OS + SS\pi)$ combined per-event mistag prediction ω.

The PDF consisted of three components, signal (S), prompt background (P), and long lived background (L). The mass PDF of the signal component consisted of a single Gaussian. We assumed both background components had similar mass distributions (an exponential with a single shape parameter). The proper time PDF of the signal component can be written as

$\mathcal{P}_S(t,d,\omega) = \mathcal{P}_S(t,d|\omega)\cdot\mathcal{P}_S(\omega)$. The first term is a conditional PDF as it depends on the value of ω, the second term describes the distribution of ω. The background parameterization of the decay time factorizes, $\mathcal{P}_B(t,d,\omega) = \mathcal{P}_B(t,d)\cdot\mathcal{P}_B(\omega)$.

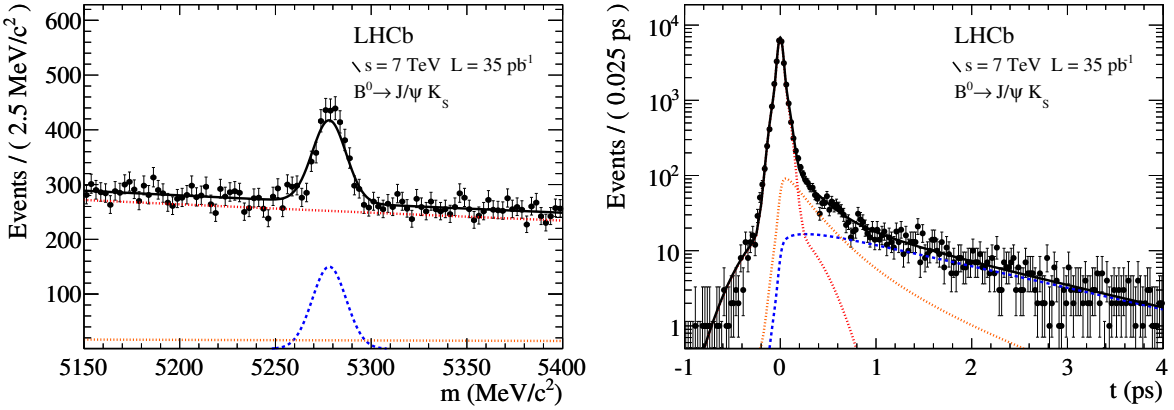


FIGURE B.3: Reconstructed mass (left) and proper time (right) distributions of $B^0 \to J/\psi K_S$ candidates. Overlayed are projections of the component PDF's used in the fit: full PDF (solid black), signal (dashed blue), prompt background (dotted red), long lived background (dotted orange).

The conditional decay time PDF of the $B^0 \to J/\psi K_S$ signal is given by $\mathcal{P}_S(t,d|\omega) = \epsilon_s(t)\cdot\Big(\mathcal{P}_{S,CP}(t,d|\omega)\otimes\mathcal{R}(t)\Big)$. Here, the PDF $\mathcal{P}_{S,CP}(t,d|\omega)$ describes decay, mixing and CP violation in the B^0_d system, and thus depends on the CP violation parameters $S_{J\psi K^0_S}$ and $C_{J\psi K^0_S}$. It was convoluted with a triple Gaussian resolution function $\mathcal{R}(t)$ which was chosen equal between all subsamples of the simultaneous fit. The efficiency function $\epsilon_s(t)$ describes acceptance effects observed in the biased subsample at low proper times.

The decay time PDF of the prompt background was just $\mathcal{R}(t)$. That of the long lived background in the unbiased samples was the sum of two exponentials with different pseudo lifetimes, in the biased samples they were simplified to a single exponential. These exponentials were also convoluted with the proper time resolution $\mathcal{R}(t)$, in order to properly describe the drop-off at $t < 0$.

In the fit to the $B^0 \to J/\psi K_S$ channel we fixed the mixing frequency Δm_d to its nominal value of $\Delta m_d = (0.507 \pm 0.005)$ ps^{-1} [6], and $C_{J\psi K^0_S} = 0$. In total, there were 27 floating parameters: the CP parameter $S_{J\psi K^0_S}$, the B^0_d lifetime, the B^0_d mass, twelve event yields, four parameters of the long-lived proper time background, five parameters of the time resolution, the mass signal resolution, and two parameters of the mass background shape.

We considered several sources of systematic uncertainty, possibly affecting the measurement of $S_{J\psi K^0_S}$. They are summarized in Table B.1. The leading contribution arises from the uncertainty in the tagger calibration. This is expressed through the uncertainty of the calibration parameters p_0, p_1. We propagated their uncertainties on the calibration (multiplying gaussian priors to the full pdf). Also the sample was refitted with the parameters obtained with the control channel

$B^0 \to J/\psi K^*$. The measurement was clearly statistically limited, we therefore, estimated the systematic errors in a conservative fashion.

Source	uncertainty
tagger calibration	0.067
per-event mistags p.d.f.	0.012
Δm_d uncertainty, z scale	0.0017
proper time resolution	0.0085
biased events acceptance	0.0042
production asymmetry	0.024
total (sum in squares)	0.073

TABLE B.1: Systematic uncertainties to $S_{J\psi K^0_S}$ in absolute terms.

The result of the maximum likelihood fit to the full data sample gave the first measurement of $\sin 2\beta$ at LHCb. The fit to the asymmetry can be seen in Figure B.4. The measured value was

$$S_{J\psi K^0_S} = 0.53^{+0.28}_{-0.29}(\text{stat}) \pm 0.08(\text{syst}) \tag{B.5}$$

This result was compatible with the world average ($\sin 2\beta = 0.673 \pm 0.023$ [6]) and dominated by the statistical uncertainty. The calculated statistical significance of a non-zero CP violation was 1.8. This constitutes a possible hint of an evidence for CP violation in the golden channel $B^0 \to J/\psi K_S$, for the first time in LHCb.

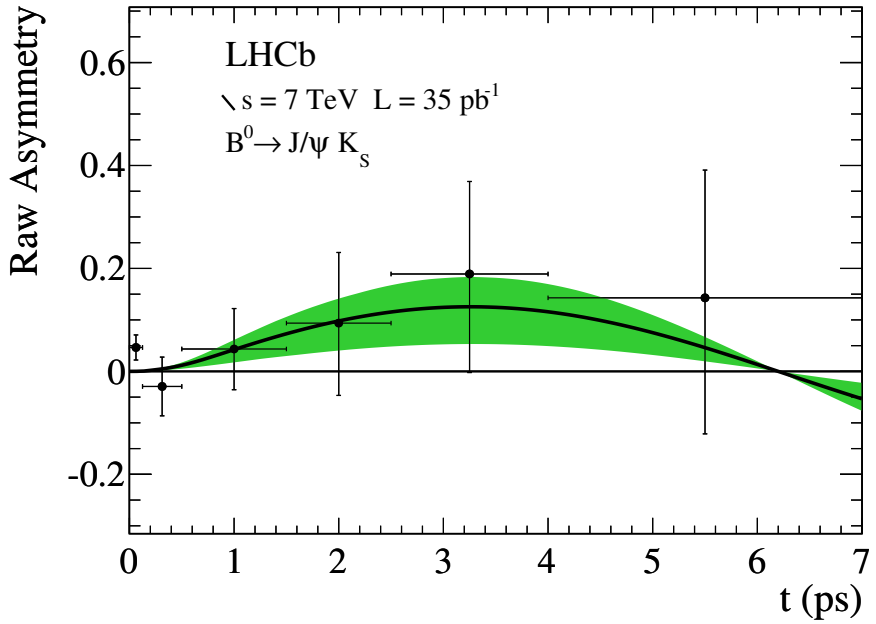


FIGURE B.4: Time dependent raw CP asymmetry in $B^0 \to J/\psi K_S$. The solid curve is the full p.d.f. (signal and background) overlayed to the data points. The green band corresponds to the one standard deviation statistical error.

The nominal fit without the Standard Model constraint $C_{J\psi K^0_S} = 0$ was also performed. In this case, we found $C_{J\psi K^0_S} = 0.28^{+0.32}_{-0.32}$, $S_{J\psi K^0_S} = 0.38^{+0.34}_{-0.35}$, quoting statistical errors only. The correlation between both parameters was $\rho(S_{J\psi K^0_S}, C_{J\psi K^0_S}) = 0.53$. Results were compatible with world average measurements.

B.1.3 Measurement of the B^0_s - $\bar{B}^0_s$ oscillation frequency Δm_s

After the observation of B^0 - $\bar{B}^0$ mixing and the measurement of its strength in 1987 [111], it took further 19 years for the B^0_s - $\bar{B}^0_s$ frequency to be measured for the first time [112],[113]. This is mainly due to the fact that the B^0_s - $\bar{B}^0_s$ oscillation frequency is 35 times larger than that for the B^0 - $\bar{B}^0$ system, posing a considerable challenge for the decay time resolution of detectors. The oscillation frequency in the B^0_s - $\bar{B}^0_s$ system, Δm_s, is given by the mass difference between the heavy and light mass eigenstates. For the LHCb experiment, the ability to resolve these fast B^0_s - $\bar{B}^0_s$ oscillations is a prerequisite for many physics analyses (like for the study of the time-dependent CP asymmetry of $B^0_s \to J/\psi\phi$ decays [28]).

The world average measurement previous to the LHCb measurement was $\Delta m_s = 17.77 \pm 0.10 \pm 0.07$ ps^{-1} [6]. The first measurement of the B^0_s - $\bar{B}^0_s$ oscillation frequency Δm_s at LHCb experiment has been measured performed using $B^0_s \to D^-_s \pi^+$ and $B^0_s \to D^-_s \pi^+ \pi^- \pi^+$ decays collected on 36 pb^{-1} of pp collisions at $\sqrt{s} = 7$ TeV in the 2010 physics run [37]. For this analysis, a total of 1381 $B^0_s \to D^-_s \pi^+$ and $B^0_s \to D^-_s \pi^+ \pi^- \pi^+$ signal decays were reconstructed, with average decay time resolutions of 44 and 36 fs respectively.

This analysis used B^0_s candidates reconstructed in four flavour-specific decay modes,[1] namely $B^0_s \to D^-_s(\phi(K^+K^-)\pi^-)\pi^+$, $B^0_s \to D^-_s(K^{*0}(K^+\pi^-)K^-)\pi^+$, $B^0_s \to D^-_s(K^+K^-\pi^-)\pi^+$and $B^0_s \to D^-_s(K^+K^-\pi^-)\pi^+\pi^-\pi^+$. Thus the flavour of the B^0_s at the time of its decay is given by the charges of the final state particles of the decay.

The offline event selection criteria were optimized individually for each of the four decay modes, exploiting the long B^0_s lifetime, the p and p_T of the B^0_s candidate and of its decay products. Particle identification variables and track and vertex quality were used too. Finally, cuts on the impact parameter significance of the reconstructed D^-_s and its distance of closest approach to the primary vertex were applied together with requiring the reconstructed mass is consistent with the PDG value [6]. The full selection is detailed in [37].

An unbinned likelihood method is employed to fit simultaneously the invariant mass and decay time distributions of the four decay modes. The probability density functions (PDFs) for the signal and for the background in each of the four modes can be written as

[1] inclusion of charge-conjugated modes is implied.

$$\mathcal{P} = \mathcal{P}_m(m)\,\mathcal{P}_t(t,q|\sigma_t,\eta)\,\mathcal{P}_{\sigma_t}(\sigma_t)\,\mathcal{P}_\eta(\eta), \tag{B.6}$$

where m is the reconstructed invariant mass of the B_s^0 candidate, t is its reconstructed decay time and σ_t is the event-by-event estimate of the decay time resolution given by the event reconstruction algorithm. The tagging decision q can be 0 (no tag), -1 (different flavour at production and decay) or $+1$ (same flavour at production and decay). The predicted event-by-event mistag probability η can take values between 0 and 0.5.

The $\mathcal{P}_m$ term describe the invariant mass distribution, which is fitted simultaneously in the four decay modes taking into account contributions from signal, combinatorial background and b decays backgrounds. The signals are described by Gaussian distributions. The fit constrains the mean of the Gaussian distributions to be the same for all four decay modes, whereas it allows the width to be different for the $B_s^0 \to D_s^-\pi^+$ and the $B_s^0 \to D_s^-\pi^+\pi^-\pi^+$modes, respectively. The combinatorial backgrounds are described by exponential functions. The b decay backgrounds include partially reconstructed B_s^0 decays, as well as fully and partially reconstructed B_d^0 and Λ_b decays with one mis-identified daughter particle. In Fig B.5 it is shown the invariant mass spectra for the four decay modes after all selection criteria.

$\mathcal{P}_t$ is a conditional probability depending on σ_t and η which describes the decay time distribution. The terms $\mathcal{P}_{\sigma_t}$ and $\mathcal{P}_\eta$ are required to ensure the proper relative normalization of $\mathcal{P}_t$ for signal and background [114]. The distribution of the decay time t of the signal is described by

$$\begin{aligned}\mathcal{P}_t(t,q|\sigma_t,\eta) \quad \propto \quad & \left\{\Gamma_s e^{-\Gamma_s t}\,\frac{1}{2}\left[\cosh(\frac{\Delta\Gamma_s}{2}t) + q[1-2\omega(\eta)]\cos(\Delta m_s t)\right]\theta(t)\right\} \\ & \otimes\, G(t,\sigma_t)\,\epsilon(t)\,\epsilon_{\rm s}.\end{aligned} \tag{B.7}$$

where Γ_s is the B_s^0 decay width and $\Delta\Gamma_s$ the decay width difference between the heavy and the light mass eigenstates. In the fit, $\Delta\Gamma_s$ is fixed to its PDG value of $0.09\,\Gamma_s$ [6]. The step function $\theta(t)$ restricts the PDF to positive decay times. The true decay time is convolved with a gaussian decay time resolution (G). The shape of the decay time distribution is distorted by trigger and offline seletion, this is accounted by introducing an acceptance function $\epsilon(t)$. The decay time distributions for the b decay backgrounds from B_d^0 and Λ_b decays are described in the same way as that for signal B_s^0 candidates, using the PDG values for their lifetimes and $\Delta\Gamma = 0$. The shape of the decay time distribution for the combinatorial background is described by the sum of two exponential functions multiplied by a second order polynomial. Within its statistical uncertainty the reconstructed B_s^0 lifetime agrees with the PDG value [6].

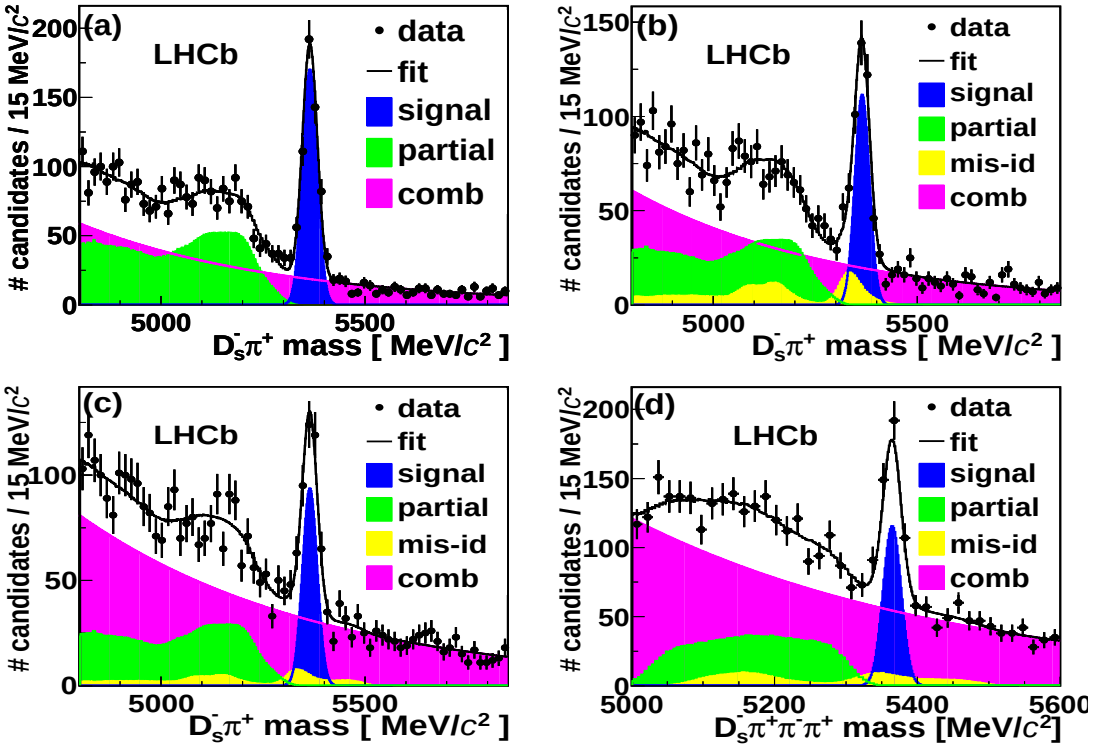


FIGURE B.5: Mass distributions for a) $B_s^0 \to D_s^-(\phi\pi^-)\pi^+$, b) $B_s^0 \to D_s^-(K^{*0}K^-)\pi^+$ c) $B_s^0 \to D_s^-(K^+K^-\pi^-)\pi^+$ and d) $B_s^0 \to D_s^-\pi^+\pi^-\pi^+$ candidates. The fits and the various background components are described in the text. "Partial" refers to background from partially reconstructed B_s^0 decays, "mis-id" refers to background from fully or partially reconstructed B^0 and Λ_b decays with one mis-identified daughter particle, and "comb" refers to combinatorial background.

The decay time PDF for untagged ($q = 0$) signal events is given by Eq. (B.1.3) multiplied by a factor $(1-\epsilon_s)$ instead of ϵ_s. The statistical power of the tagging is determined by the "effective" tagging efficiency for signal events and is defined as

$$\epsilon_{\text{eff}} = \epsilon_s \times \frac{1}{\sum_i W_i} \sum_i (1 - 2\omega(\eta_i))^2 \times W_i, \tag{B.8}$$

where the signal tagging efficiency ϵ_s is a free parameter in the fit of the oscillation frequency described in the next section. W_i is the probability for being a signal event as determined by the invariant mass and decay time PDFs. The index i runs over all B_s^0 candidates.

The tagging algorithms were optimized and calibrated using $B \to D^{*-}\mu^+X$ and $B^+ \to J/\psi K^+$ events 4. As trigger and selection cuts can bias the distributions of the event properties used by the tagging algorithms, this could result in a biased estimation for the $B_s^0 \to D_s^-\pi^+$ and $B_s^0 \to D_s^-\pi^+\pi^-\pi^+$ events. Therefore, a re-calibration was performed using a sample of 6,000 $B^0 \to D^-\pi^+$ events, which have a similar topology to the $B_s^0 \to D_s^-\pi^+$ and $B_s^0 \to D_s^-\pi^+\pi^-\pi^+$ events, and were collected using the same trigger and similar selection cuts. The calibration

parameters were obtained from the maximum likelihood fit of the $B^0\bar{B}^0$ oscillation signal and found to be consistent with the original calibration.

The fit for the oscillation frequency Δm_s is performed simultaneously to all four B_s^0 decay modes and gives $\Delta m_s = 17.63 \pm 0.11$ ps^{-1} (statistical uncertainty only). Signal tagging efficiencies of $\epsilon_{\rm s} = (23.6 \pm 1.3)\,\%$ and $\epsilon_{\rm s} = (17.6 \pm 3.2)\,\%$ are found for the $B_s^0 \to D_s^-\pi^+$ and $B_s^0 \to D_s^-\pi^+\pi^-\pi^+$ modes, respectively. The combined effective tagging efficiency for all four modes is $\epsilon_{\rm eff} = (3.8 \pm 2.1)\,\%$. The likelihood profile as a function of the assumed oscillation frequency Δm_s is shown in Figure B.6 (left). The statistical significance of the signal is evaluated to be $4.6\,\sigma$ by comparing the likelihood value at the minimum of the fit with that found in the limit $\Delta m_s = \infty$. The oscillation pattern is clearly visible when the asymmetry is plotted in bins of the decay time modulo $2\pi/\Delta m_s$ (Figure B.6 right).

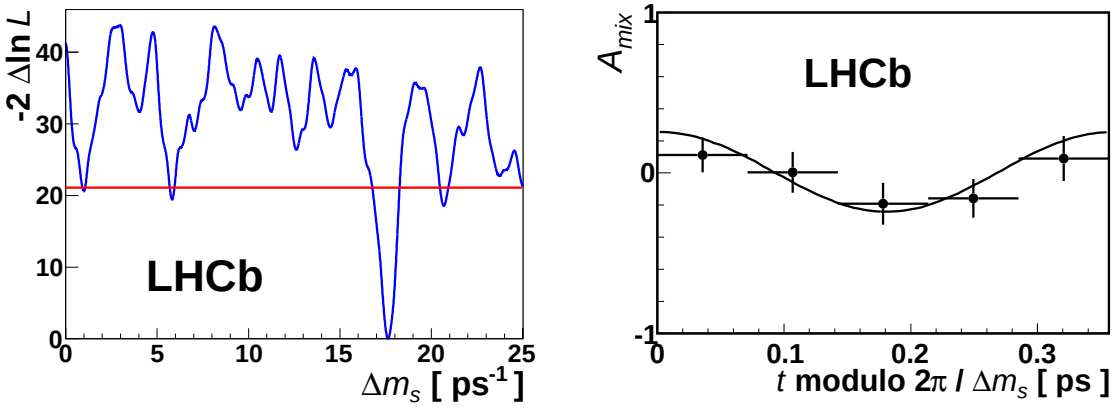


FIGURE B.6: Left: Likelihood scan for Δm_s in the range [0.0, 25.0] ps^{-1}. The line at $-2\Delta \ln L = 20.9$ indicates the value in the limit $\Delta m_s = \infty$. Right: Measured asymmetry for B_s^0 candidates in bins of the decay time t modulo $2\pi/\ \Delta m_s$. The projection of the likelihood fit is superimposed.

Various possible sources of systematic effects had been studied, the dominant source of systematic uncertainty is due to the knowledge of the absolute decay time scale of the experiment, which is dominated by the knowledge of the z scale. Other sources of systematic uncertainties are related with the description of the combinatorial background in the fit to the mass spectra, the decay time resolution and the momentum scale. Other possible sources were found to be negligible. The total systematic uncertainty is defined as the quadratic sum of the individual components and is 0.022 ps^{-1}.

The result to the fit found the following value:

$$\Delta m_s = 17.63 \pm 0.11(\text{stat}) \pm 0.02(\text{syst})\text{ps}^{-1}. \tag{B.9}$$

This was in good agreement with the previous best measurement of $\Delta m_s = 17.77 \pm 0.10(\text{stat}) \pm 0.07(\text{syst})$ ps^{-1}, reported by the CDF collaboration [113]. The results were completely dominated by statistical uncertainties and thus significant improvements are expected with larger data sets.

B.1.4 Measurement of the CP violating phase ϕ_s in the decay $B_s^0 \to J/\psi\phi$

As explained in Chapter 1, in the Standard Model (SM) CP violation arises through a single phase in the CKM quark mixing matrix [14, 22]. In neutral B meson decays to a final state which is accessible to both B and $\bar{B}$ mesons, the interference between the amplitude for the direct decay and the amplitude for decay after oscillation, leads to a time-dependent CP violating asymmetry between the decay time distributions of B and $\bar{B}$ mesons. The decay $B_s^0 \to J/\psi\phi$ allows the measurement of such an asymmetry, which can be expressed in terms of the decay width difference of the heavy (H) and light (L) B_s^0 mass eigenstates $\Delta\Gamma_s \equiv \Gamma_\text{L} - \Gamma_\text{H}$ and in a single phase ϕ_s [20]. In the SM, the decay width difference is $\Delta\Gamma_s^{SM} = 0.082 \pm 0.021$ ps^{-1} [115, 116], while the phase is predicted to be small, $\phi_s^\text{SM} = -2\arg\left(-V_{ts}V_{tb}^*/V_{cs}V_{cb}^*\right) = -0.036 \pm 0.002$ rad [27]. Contributions from physics beyond the SM could lead to much larger values of ϕ_s [117].

The first analysis of the CPV phase $\phi_s^{J/\psi}$ in the decay $B_s \to J/\psi\phi$ of the LHCb collaboration was presented in spring 2011 based on a data set of ~ 36 pb^{-1} [118]. Based on a sample of 836 $\pm$ 60 signal candidates with $t > 0.3$ ps, a two-dimensional confidence contours in the $\Delta\Gamma_s - \phi_s$ plane could be derived (Figure B.7 left). The analysis was at this time completely statistically limited. The second analysis was performed based on 370 pb^{-1} of data, which was presented at the *Lepton Photon* conference end of August 2011 and produced the best measurement ever of ϕ_s, $\Delta\Gamma_s$ and Γ_s in $B_s^0 \to J/\psi\phi$ decays, substantially improving upon previous measurements. It provided the first direct evidence for a non-zero value of $\Delta\Gamma_s$. In this last analysis, two solutions with equal likelihood were obtained, related by the transformation $(\phi_s, \Delta\Gamma_s) \mapsto (\pi - \phi_s, -\Delta\Gamma_s)$. Figure B.7 right shows the 68%, 90% and 95% CL contours in the $\Delta\Gamma_s - \phi_s$ plane.

The analysis on 370 pb^{-1} of data was performed selecting only B_s^0 candidates decaying into $J/\psi \to \mu^+\mu^-$ and $\phi \to K^+K^-$. The J/ψ and ϕ candidates are created from oppositely charged muons (with $p_T > 0.5$ GeV) and kaons, from a commmon vertex and with an invariant mass in the range 3030 - 3150 MeV for the J/ψ and within $\pm$12 MeV of the nominal ϕ mass [6]. The complete candidate selection is detailed in [104]. Only B_s^0 candidates with invariant mass (m_B) in the range of 5200-5550 MeV were considered. To supress a large fraction of prompt combinatorial background only candidates with a decay time within the range $0.3 < t < 14$ ps were kept. From a fit to the m_B distribution, shown in Figure B.8, a signal of 8492 ± 97 events was extracted.

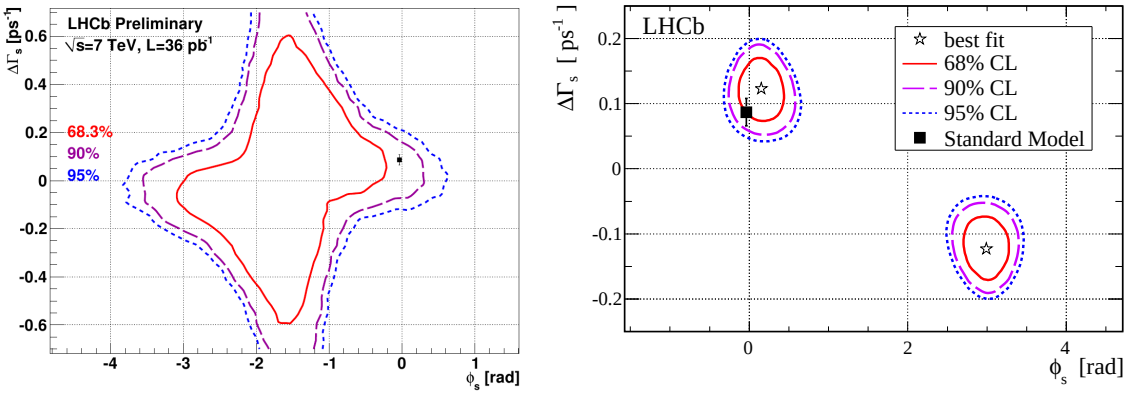


FIGURE B.7: Likelihood confidence regions in the $\Delta\Gamma_s$-ϕ_s plane. The black square and error bar corresponds to the Standard Model prediction [115, 116]. Left: from 2010 data analysis with $\sim 36 pb^{-1}$. Right: from 2011 data analysis with $\sim 370\ pb^{-1}$.

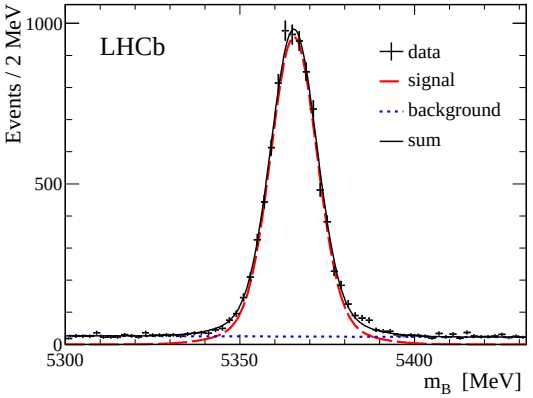


FIGURE B.8: Invariant mass distribution for $B_s^0 \to \mu^+\mu^- K^+K^-$ candidates with the mass of the $\mu^+\mu^-$ pair constrained to the nominal J/ψ mass. Curves for fitted contributions from signal (dashed), background (dotted) and their sum (solid) are overlaid.

In the $B_s^0 \to J/\psi\phi \to \mu^+\mu^- K^+K^-$ decay, the final state can be CP-even or CP-odd depending upon the relative orbital angular momentum between the J/ψ and the ϕ. In order to measure ϕ_s it is necessary to disentangle the CP-even and CP-odd components. This is achieved by analysing the distribution of the reconstructed decay angles $\Omega = (\theta, \psi, \varphi)$ in the transversity basis [119, 120].

An unbinned maximum likelihood fit was performed to the invariant mass m_B, the decay time t, and the three decay angles Ω. The probability density function (PDF) used in the fit consists of signal and background components which include detector resolution and acceptance effects. The PDF's are factorised into separate components for the mass and for the remaining observables.

The signal m_B distribution was described by two Gaussian functions with a common mean while the m_B distribution for the combinatorial background is described by an exponential function

with a slope determined by the fit.

The distribution of the signal decay time and angles is described by a sum of ten terms, each of which is the product of a time-dependent term and an angular function [120] as:

$$\frac{d^4\Gamma(B_s^0 \to J/\psi\phi)}{dt\, d\Omega} \propto \sum_{k=1}^{10} h_k(t)\, f_k(\Omega)\,. \tag{B.10}$$

The functions $h_k(t)$ are expressed in terms of ϕ_s, Γ_s, $\Delta\Gamma_s$, the B_s^0 oscillation frequency Δm_s and four complex transversity amplitudes A_i at $t = 0$. The label i takes the values $\{\perp, \|, 0\}$ for the three P-wave amplitudes and S for the S-wave amplitude. Each A_i is parametrized by its magnitude squared $|A_i|^2$ and its phase δ_i, the convention $\delta_0 = 0$ and $\sum |A_i|^2 = 1$ were adopted. The full set of expressions is given in [120–122]. Using the data in the m_B sidebands the background decay time and decay angle distributions were determined.

To account for the decay time resolution, the PDF was convoluted with a sum of three Gaussian functions with a common mean and different widths. The parameters of the resolution model are determined from a fit to the decay time distribution of prompt $J/\psi\ K^+K^-$ combinations after substracting non-J/ψ events. The decay time distribution is affected by two acceptance effects. The first one can be parametrized as $\epsilon(t) \propto (1-\beta t)$ where the factor $\beta = 0.016\ \mathrm{ps}^{-1}$ is determined from simulated events. The second, is applied exclusively by events triggered by paths that exploits large impact parameters of the decay products, leading to a drop in efficiency at small decay times, and can be described by the empirical acceptance function $\epsilon(t) \propto (at)^c\,/\,[1+(at)^c]$.

The measurement of ϕ_s requires knowledge of the flavour of the B_s^0 meson at production, for that, only the flavour specific features of the accompanying (non-signal) b-hadron decay (OS) to tag the B_s^0 flavour were exploited. The fraction of tagged events in the signal sample is $\varepsilon_{\mathrm{tag}} = (24.9\pm0.5)\%$, the effective tagging efficiency is $\varepsilon_{\mathrm{tag}}D^2 = (1.91\pm0.23)\%$. From the tagging algorithms a per-event mistag probability (ω) calibrated with data from control channels [97] is provided, the uncertainty in ω is taken into account by allowing calibration parameters described in [97] to vary in the fit with Gaussian constraints given by their estimated uncertainties. Both tagged and untagged events are used in the fit. The untagged events dominate the sensitivity to the lifetimes and amplitudes.

The sensitivity to ϕ_s is diluted by the decay time resolution and mistag probability. Systematic uncertainties from these sources and from the mixing frequency are absorbed in the statistical uncertainties. Other systematic uncertainties were taken into account.

The results of the fit for the main observables are shown in Table B.2. Figure B.9 shows the data distribution for decay time and angles with the projections of the best fit PDF overlaid. To assess the overall agreement of the PDF with the data it was calculated the goodness of fit based on the point-to-point dissimilarity test [123]. The p-value obtained was 0.68. Figure B.7

TABLE B.2: Fit results for the solution with $\Delta\Gamma_s > 0$ with statistical and systematic uncertainties.

parameter	value	$\sigma_{\text{stat.}}$	$\sigma_{\text{syst.}}$
Γ_s [ps^{-1}]	0.657	0.009	0.008
$\Delta\Gamma_s$ [ps^{-1}]	0.123	0.029	0.011
$\lvert A_\perp(0)\rvert^2$	0.237	0.015	0.012
$\lvert A_0(0)\rvert^2$	0.497	0.013	0.030
$\lvert A_{\text{S}}(0)\rvert^2$	0.042	0.015	0.018
$\delta_\perp$ [rad]	2.95	0.37	0.12
δ_{S} [rad]	2.98	0.36	0.12
ϕ_s [rad]	0.15	0.18	0.06

(right) shows the 68%, 90% and 95% CL contours in the $\Delta\Gamma_s$-ϕ_s plane. These contours are obtained from the likelihood profile after including systematic uncertainties.

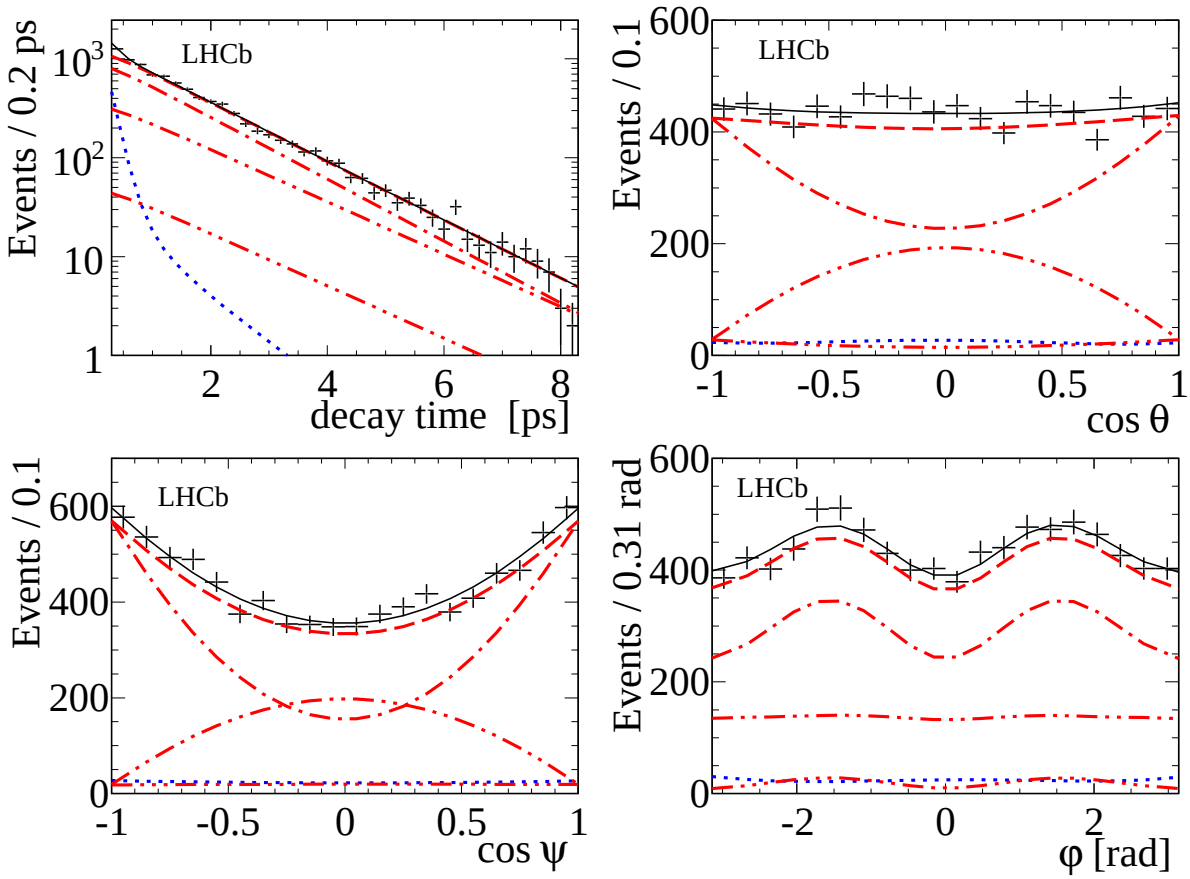


FIGURE B.9: Projections for the decay time and transversity angle distributions for events with m_B in a ± 20 MeV range around the B_s^0 mass. The points are the data. The dashed, dotted and solid lines represent the fitted contributions from signal, background and their sum. The remaining curves correspond to different contributions to the signal, namely the CP-odd P-wave (dashed with single dot), the CP-even P-wave (dashed with double dot) and the S-wave (dashed with triple dot).

In summary, in a sample of 0.37 fb^{-1} of pp collisions at $\sqrt{s} = 7$ TeV collected with the LHCb detector we observe 8492 ± 97 $B_s^0 \to J/\psi K^+K^-$ events. With these data we performed the most precise measurements of ϕ_s, $\Delta\Gamma_s$ and Γ_s in $B_s^0 \to J/\psi\phi$ decays, substantially improving upon previous measurements [124, 125] and providing the first direct evidence for a non-zero value of $\Delta\Gamma_s$. Two solutions with equal likelihood are obtained, related by the transformation $(\phi_s, \Delta\Gamma_s) \mapsto (\pi - \phi_s, -\Delta\Gamma_s)$. The solution with positive $\Delta\Gamma_s$ is

$$\begin{aligned}
\phi_s &= 0.15 \pm 0.18 \text{ (stat)} \pm 0.06 \text{ (syst)rad}, \\
\Gamma_s &= 0.657 \pm 0.009 \text{ (stat)} \pm 0.008 \text{ (syst)ps}^{-1}, \\
\Delta\Gamma_s &= 0.123 \pm 0.029 \text{ (stat)} \pm 0.011 \text{ (syst)ps}^{-1}
\end{aligned}$$

and is in agreement with the Standard Model prediction [115, 116]. In Figure B.10 a comparison of the B_s^0 mixing results with data available in summer 2011 is shown.

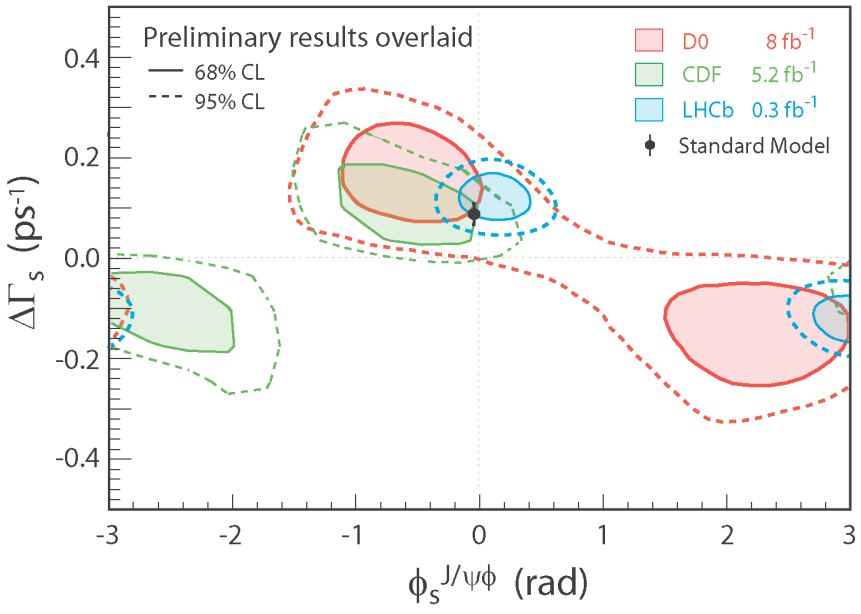


FIGURE B.10: A comparison of the CDF [124], D0 [125] and LHCb [122] collaboration B_s^0 mixing results.

Bibliography

[1] S. L. Glashow, *Partial symmetries of weak interactions*, Nucl. Phys. **22** (1961) 579-588.

[2] S. Weinberg, *A model of leptons*, Phys. Rev. Lett. **19** (1967) 1264-1266.

[3] A. Salam, *Weak and electromagnetic interactios*, Proc. of the 8th Nobel Symposium on 'Elementary particle theory, relativistic groups and analyticity', Stockholm, Sweden, 1968, edited by N. Svartholm, 367-377.

[4] M. B. Gavela et al., *Standard Model CP-Violation and Baryon asymmetry*, Mod. Phys. Lett. **9A** (1994) 795.

[5] ALEPH Collaboration, *Determination of the Number of Light Neutrino Species*, CERN EP 89-132, Oct 1989. 15pp, Phys. Lett. B **231** (1989) 519.

[6] K. Nakamura et al., (Particle Data Group), *Review of particle physics*, J. Phys. G 37, 075021 (2010).

[7] D. J. Gross and F. Wilczek, *Ultraviolet behaviour of non-abelian gauge theories*, Phys. Rev. Lett. **30 (26)** (1973) 1343-1346.

[8] H. D. Politzer, *Reliable Perturbative Results for Strong Interactions*, Phys. Rev. Letters **30** (1973) 1346.

[9] D. J. Gross and F. Wilczek, *Asymptotically Free Gauge Theories.*, Phys. Rev. **D8** (1973) 3633.

[10] H. D. Politzer, *Asymptotic Freedom: An Approach to Strong Interactions*, Phys. Rep. **14** (1974) 129.

[11] http : //www.nobelprize.org/nobel_prizes/physics/laureates/2004.

[12] S. Bethke, α_s *2002*, arXiv:hep-ex/0211012v1 (2002).

[13] P. Renton, *Electroweak Interactions*, Cambridge University Press. 1990.

[14] N. Cabibbo, *Unitarity Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (1963) 531-533.

[15] C. S. Wu et al., *Experimental Test of Parity Conservation in Beta Decay*, Phys. Rev. Lett. **105** (1957) 1413-1415.

[16] J. Cronin et al., *Evidence for the 2π Decay of the K_2^0 Meson*, Phys. Rev. Lett. **13** (1964) 138.

[17] BaBar Collaboration, *Observation of CP violation in the B^0 meson system*, Phys. Rev. Lett. **87** (2001) 091801.

[18] Belle Collaboration, *Observation of large CP violation in the neutral B meson system*, Phys. Rev. Lett. **87** (2001) 091802.

[19] LHCb Collaboration, *Evidence for CP violation in time-integrated $D^0 \to h^- h^+$ decay rates*, arXiv:1112-0938

[20] I. I. Bigi and A. I. Sanda, *Notes on the observability of CP violations in B decays* Nucl. Phys. B **193** (1981) 85.

[21] M. E. Peskin and D. V. Schroeder, *An introduction to quantum field theory* Addison-Wesley Advanced Book Program. 1995.

[22] M. Kobayashi and T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, Progress of Theoretical Physics **49**, 652 (1973).

[23] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Let. **51** (1983)

[24] I. I. Bigi and A. I. Sanda, *CP Violation*, Cambridge University Press. 2000.

[25] The CKM Fitter group, http : //ckmfitter.in2p3.fr.

[26] The UT Fitter group, http : //utfit.org.

[27] J. Charles et al., *CP Violation and the CKM Matrix: Assessing the Impact of the Asymmetric B Factories*, The European Physics Journal C, **41**, 1 (2005), arXiv:hep-ph/0406184

[28] The LHCb Collaboraion, *Measurement of the CP violating phase ϕ_s in the decay $B_s^0 \to J/\psi\phi$*, arXiv:1112.3183; CERN-PH-EP-2011-214; LHCB-PAPER-2011-021.

[29] S. W. Herb et al., *Observation of a Dimuon Resonance at 9.5* GeV*in 400* GeV*Proton-Nucleus Collisions*, Phys. Rev. Lett. **39** (1977) 252.

[30] W. R. Innes et al., *Observation of Structure in the Υ Region*, Phys. Rev. Lett. **39** (1977) 1240.

[31] E. Fernandez et al., *Lifetime of Particles Containing b Quarks*, Phys. Rev. Lett. **51** (1983) 1022.

[32] N. Lockyer et al., *Measurement of the Lifetime of Bottom Hadrons*, Phys. Rev. Lett. **51** (1983) 1316.

[33] UA1 Collaboration, *Search for $B^0\bar{B}^0$ oscillations at the CERN proton-antiproton collider*, Phys. Lett. **B186** (1987) 247.

[34] ARGUS Collaboration, *Observation of $B^0\bar{B}^0$ mixing*, Phys. Lett. **B192** (1987) 245.

[35] G. C. Branco, L. Lavoura and J. P. Silva, *CP Violation*, Oxford Science Publications. 1999.

[36] The LHCb Collaboration, *Measurement of Δm_d in $B^0 \to D^-(K^+\pi^-\pi^-)\pi^+$*, LHCb-CONF-2011-010 (2011)

[37] The LHCb Collaboration, *Measurement of Δm_s in the decay $B_s^0 \to D_s^-(K^+K^-\pi^-)(3)\pi$*, LHCb-CONF-2011-005 (2011)

[38] The LHCb Collaboration, *Measurement of direct CP violation in charmless charged two-body B decays at LHCb*, LHCb-CONF-2011-011, **27** (2011)

[39] The Belle Collaboration et al., Phys. Rev. Lett. D **77**, 031802 (2008)

[40] The BaBar Collaboration et al., Phys. Rev. Lett. D **79**, 072009 (2009)

[41] The HFAG Collaboration et al., arXiv:0808.1297

[42] The LHCb Collaboration, *Measurement of CP violation in the time dependent analysis of $B^0 \to J/\psi K_S$ decays with 2010 data*, LHCb-CONF-2011-004, (2011)

[43] Z. Ligeti et al., *Implications of the CP asymmetry in semileptonic B decay*, arXiv:hep-ph/0604112 (2006).

[44] S. P. Martin, *A supersymmetry primer*, arXiv:hep-ph/9709356v5 (2008).

[45] S. Dimopoulos, S. Raby and F. Wilczek, *Supersymmetry and the scale of the unification*, Phys. Rev. Lett D, **24(6)** (1981) 1681-1683.

[46] Arkani-Hamed, S. Dimopoulos and G. Dvali, *The Hierarchy problem and new dimensions at a millimeter*, arXiv:hep-ph/9803315 (1998).

[47] L. Randall and R. Sundrum, *Large Mass Hierarchy from a Small Extra Dimension.* arXiv:hep-ph/9905221 (1999).

[48] A. Martin, *Technicolor signals at the LHC*, arXiv:hep-ph/0812.1841v2 (2008).

[49] The LHCb Collaboration, S. Amato et al., *A large Hadron Collider Beauty experiment for precision measurements of CP-violation and rare decays. The LHCb Technical Proposal*, CERN/LHCC 98-4.

[50] http : //home.thep.lu.se/ torbjorn/Pythia.html

[51] The LHCb Collaboration, *Measurement of* $\sigma(pp \to b\bar{b}X)$ *at* $\sqrt{s} = 7$ TeV *in the forward region*, Phys. Lett. B **694** (Sep, 2010) 209-216, arXiv:hep-ex/1009.2731.

[52] The LHCb Collaboration, *The LHCb detector at the LHC*, JINST 3 (2008) S08005.

[53] The LHCb Collaboration, *The LHCb Reoptimized Detector and Performances*, CERN/LHCC 2003-030, September 2003.

[54] The LHCb Collaboration, *LHCb Vertex Locator Technical Design Report*, CERN/LHCC 2001-0011, May 2001.

[55] The LHCb Collaboration, *LHCb Magnet Technical Design Report*, CERN/LHCC 2000-007, January 2000.

[56] The LHCb Collaboration, *LHCb Inner Tracker Technical Design Report*, CERN/LHCC 2002-029, November 2009.

[57] The LHCb Collaboration, *LHCb Outer Tracker Technical Design Report*, CERN/LHCC 2001-024, September 2001.

[58] G.W van Apeldoorn et al., *Beam Test of Final Modules and Elecgtronics of the LHCb Outer Tracker in 2005*, Technical Report LHCb-2005-076, 2005.

[59] The LHCb Collaboration, *LHCb RICH Technical Design Report*, CERN/LHCC 2000-037, September 2000.

[60] The LHCb Collaboration, *LHCb Calorimeters Technical Design Report*, CERN/LHCC 2000-036, September 2000.

[61] E. Guschin et al., *Design and Construction of the LHCb scintillator PAD/preshower Detector*, LHCb Note 2001-138, December 2001.

[62] M. Calvo, *Backsplash Effects on SPD and Flavour Tagging in LHCb*, CERN-Tesis 2006-011, January 2006.

[63] M. Grabalosa, E. Graugés, H. Ruiz et al., *SPD Calibration*, LHCb 2007-139, October 2007.

[64] R. Vazquez, M. Calvo, D. Esperante et al., *Calibration and aligment of the SPD detector with cosmic rays and first LHC collisions*, LHCb-PUB-2011-024, July 2011.

[65] The LHCb Collaboration, *LHCb Muon System Technical Design Report*, CERN/LHCC 2001-010, May 2001.

[66] The LHCb Collaboration, *LHCb Trigger System Technical Design Report*, CERN/LHCC 2003-031, September 2003.

[67] The LHCb Collaboration, *LHCb Online system, data acquisition and experiment control Technical Design Report*, CERN/LHCC 2002-011, April 2002.

[68] The LHCb Collaboration, *LHCb Computing Technical Design Report*, CERN/LHCC 2005-019, June 2005.

[69] The LHCb Collaboration, *The* GAUSS *Project*, http : //lhcb − release − area.web.cern.ch/LHCb − release − area/DOC/gauss/.

[70] The LHCb Collaboration, *The* BOOLE *Project*, http : //lhcb − release − area.web.cern.ch/LHCb − release − area/DOC/boole/.

[71] The LHCb Collaboration, *The* MOORE *Project*, http : //lhcb − release − area.web.cern.ch/LHCb − release − area/DOC/moore/.

[72] The LHCb Collaboration, *The* BRUNEL *Project*, http : //lhcb − release − area.web.cern.ch/LHCb − release − area/DOC/brunel/.

[73] The LHCb Collaboration, *The* DAVINCI *Project*, http : //lhcb − release − area.web.cern.ch/LHCb − release − area/DOC/davinci/.

[74] The LHCb Collaboration, *The* PANORAMIX *Project*, http : //lhcb − release − area.web.cern.ch/LHCb − release − area/DOC/panoramix/.

[75] WLCG, The WLCG homepage, http : //lcg.web.cern.ch/lcg/.

[76] The Ganga website: http : //ganga.web.cern.ch/ganga/

[77] M. Calvi, O. Dormond and M. Musy. *LHCb Flavour Tagging Performance*, LHCb-2003-115.

[78] M. Calvi, O. Leroy and M. Musy. *Flavour Tagging Algorithms and Performances in LHCb*, LHCb-2007-058.

[79] E. Polycarpo and M. Gandelman. *The Performance of the LHCb Muon Identification Procedure*, LHCb 2007-145.

[80] M. Needham. *Clone Track Identification using the Kullback-Liebler Distance*, LHCb-2008-002. CERN-LHCb-2008-002. LPHE-2008-002.

[81] http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/davinci/releases/latest/

[82] http : //root.cern.ch/root/html/TMultiLayerPerceptron.html

[83] http : //tmva.sourceforge.net

[84] L. Garrido and V. Gaitan *Use of Neural Nets to measure the τ polarization and its bayesian interpretation*, International Journal of Neural Systems, Vol. 2, No. 3 (1991).

[85] J. Babel. *Analyse d'un scénario de couplages sZb non standard : Étiquetage des saveurs auprès de l'expérience LHCb*, PhD thesis, Aix-Marseille University, CERN-THESIS-2007-017 (2006).

[86] T.Brambach et al., *Optimization and Calibration of the Tagging performances using 2010 data*, LHCb-ANA-2011-003.

[87] M. Calvo, M. Grabalosa and M. Musy, *Combination of same-side with opposite-side flavour tagging*, LHCb-PUB-2009-027.

[88] M. Calvi et al., *Study of $B^0 \to D^{*-}\mu^+\nu_\mu$ decay in LHCb and measurement of Flavour Tagging Performances*, LHCb Internal Note in preparation.

[89] M. Calvi et al., *Calibration of flavour tagging with $B^+ \to J/\psi K^+$ and $B^0 \to J/\psi K^*$ control channels at LHCb*, LHCb-2009-020.

[90] T. Brambach et al., *Measurement of the B^0 mixing frequency Δm_d in $B^0 \to J/\psi K^*$ and $B^0 \to D^-\pi^+$*, LHCb-2012-018.

[91] The LHCb Collaboration, *Measurement of Δm_s in the decay $B_s \to D_s(K^+K^-\pi^-)\pi^+$ using opposite-side and same-side flavour tagging algorithms*, LHCb-CONF-2011-050

[92] Kazu Akiba et al., *Measurement of Δm_s in the decay $B_s \to D_s(K^+K^-\pi^-)\pi^+$ and $B_s \to D_s(K^+K^-\pi^-)\pi^+\pi^+\pi^-$*, LHCb-ANA-2011-005

[93] M. Pivk and F.R. Le Diberder, Nucl. Inst. Meth.A 555, 356-369, 2005.

[94] V. Fave, M. Calvo and P. Szczypka. *Optimization of Same Side Kaon Tagger with 2010 data*, LHCb-INT-2011-032.

[95] M. Calvi et al., *Optimization and Calibration of the Same Side Kaon tagging algorithm using hadonic B_s^0 decays in 2011 data*, LHCb-ANA-2011-103.

[96] The LHCb Collaboration, *Opposite-side flavour tagging of B mesons at the LHCb experiment*, LHCb-PAPER-2011-027, CERN-PH-EP-2012-039. Submitted to Eur. Phys. J. C arXiv:1202.4979.

[97] T. Brambach et al., *Opposite side and same side pion tagging performances using 2011 data*, LHCb-ANA-2012-005.

[98] R.Aaij et al., *Tagged time-dependent angular analysis of $B_s^0 \to J/\psi\phi$ decays with $337pb^{-1}$ at LHCb*, LHCb-ANA-2011-036.

[99] *B decays at LHC*, hep-ph/0003238.

[100] M. Calvi, *Selection of $B^0 \to D^{*-}\mu^+\nu_\mu$ Events and Flavour Tagging Studies*, LHCb 2007-036.

[101] G. Cowan et al., *Measurements of B^0 mesons production cross-section in pp collisions at $\sqrt{s} = 7$ TeV using $B^0 \to D^{*-}\mu^+\nu_\mu + X$ decays*, LHCb-ANA-2010-008.

[102] LHCb Collaboration, *Measurements of B^0 mesons production cross-section in pp collisions at $\sqrt{s} = 7$ TeV using $B^0 \to D^{*-}\mu^+\nu_\mu + X$ decays*, LHCb-CONF-2010-012.

[103] W. Verkerke and D. P. Kirkby, in the Proceedings of 2003 Conference for Computing in High-Energy and Nuclear Physics (CHEP 03), La Jolla, California, 24-28 Mar 2003, arXiv:physics/0306116

[104] R. Aaij et al., *Selections and lifetime measurements for exclusive $b \to J/\psi X$ decays with $J/\psi \to \mu\mu$ with 2010 data*, LHCb-ANA-2011-001.

[105] R. Aaij et al., *Tagged time-dependent angular analysis of $B_s^0 \to J/\psi\phi$ decays with $1.03fb^{-1}$*. LHCb-ANA-2012-004.

[106] The HFAG collaboration, E. Barbiero et al., arXiv:0808.1297

[107] M. Grabalosa for the LHCb collaboration, *First measurements of $\sin 2\beta$ at LHCb with $B^0 \to J/\psi K_S$*, PoS, The 13th International Conference on B−Physics Hadron Machines, Beauty 2011. 4 - 8, 2011.

[108] T. Brambach et al., *Measurement of CP violation in the time-dependent analysis of $B^0 \to J/\psi K^0$ decays with the 2010 data*, LHCb-ANA-2011-004.

[109] The LHCb collaboration, *Selections and lifetime measurements for exclusive $b \to J/\psi$ X decays with $J/\psi \to \mu^+\mu^-$ with 2010 data*, LHCb-CONF-2011-001

[110] The LHCb collaboration, *Optimization and calibration of the LHCb flavour tagging performance using 2010 data*, LHCb-CONF-2011-003

[111] ARGUS collaboration, H. Albrecht et al., *Observation of B_d^0-$\bar{B}_d^0$ mixing*, Phys. Lett. B 192, 245 (1987).

[112] D0 collaboration, V. M. Abazov et al., *Direct Limits on the B_s^0 Oscillation Frequency*, Phys. Rev. Lett. **97** 021802 (2006).

[113] CDF collaboration, A. Abulencia et al., *Measurement of the B_s^0- barB_s^0 Oscillation Frequency*, Phys. Rev. Lett. **97** 062003 (2006).

[114] G. Punzi, *Comments on likelihood fits with variable resolution*, eConf C030908 (2003) WELT002.

[115] A. Lenz and U. Nierste, *Theoretical update of $B_s - \bar{B}_s$ mixing*, JHEP 0706 (2007) 072. arXiv:hep-ph/0612167.

[116] A. Lenz and U. Nierste, *Numerical updates of lifetimes and mixing parameters of B mesons*, arXiv:1102.4274 [hep-ph].

[117] A. J. Buras, PoS EPS-HEP2009 (2009) 024, arXiv:0910.1032.

[118] The LHCb Collaboration, *Tagged time-dependent angular analysis of $B_s^0 \to J/\psi\phi$ decays with the 2010 LHCb data*, LHCb-CONF-2011-006.

[119] A. S. Dighe, I. Dunietz et al., *Angular distributions and lifetime differences in $B_s \to J/\psi\phi$ decays*, Phys. Lett. **B369**, 144-150 (1996). arXiv:hep-ph/9511363.

[120] A. S. Dighe, I. Dunietz, R. Fleischer, *Extracting CKM phases and $B_s - \bar{B}_s$ mixing parameters from angular distributions of nonleptonic B decays*, Eur. Phys. J. **C6**, 647-662 (1999). arXiv:hep-ph/9804253.

[121] Y. Xie, P. Clarke et al., *Determination of 2beta(s) in B0(s) —¿ J/psi K+ K- Decays in the Presence of a K+ K- S-Wave Contribution*, JHEP **0909** (2009) 074. arXiv:0908.3627 [hep-ph].

[122] The LHCb Collaboration, *Tagged time-dependent angular analysis of $B_s^0 \to J/\psi\phi$ decays with 337 pb^{-1} at LHCb*, LHCb-CONF-2011-049.

[123] M. Williams, *How good are your fits? Unbinned multivariate goodness-of-fit tests in high energy physics*, JINST **5** (2010) P09004. arXiv:1006.3019 [hep-ex].

[124] The CDF Collaboration, public note CDF/ANAL/BOTTOM/PUBLIC/10206 (2010).

[125] The D0 Collaboration, D0 Conference note 6098-CONF (2010).